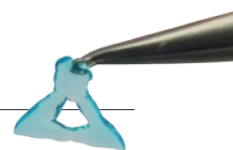


THIS WEEK



EDITORIALS

NEUROSCIENCE Searching for the seat of pain in the brain **p.456**

WORLD VIEW Islamic scholars must focus on political groups **p.457**

ORGANS Gel with cells and nutrients could print tissues **p.458**

When right beats might

The final act in a long-running Italian saga should bring tighter controls on unproven stem-cell therapies, both at home and abroad.

What impact can scientists have on the murky world of politics? When the Italian stem-cell researcher Elena Cattaneo was appointed senator in her nation's parliament in 2013, she carried the hopes of her colleagues that she could make a difference. Italian courts and politicians were authorizing unproven and potentially dangerous stem-cell therapies. The emotions of vulnerable and desperately ill people were being played with for commercial gain. Reporters, including those from this journal, who tried to expose the shameful situation were lied to, obstructed and threatened with legal action. Signora Cattaneo went to Rome. Together with other stem-cell researchers, she helped to bring an end to the whole sorry affair. Brava!

In her research at the University of Milan, Cattaneo investigates the conversion of embryonic stem cells into mature nerve cells, and how this might one day be used to treat neurological diseases. Across the world, the development of such treatments, and how they should be tested, introduced and regulated, is pitting careful, evidence-based approaches against medical opportunists. Until recently, Italy was an example of how to get it wrong. Now it has an opportunity to show how it — and the rest of the world — can get it right.

For seven years, the Stamina Foundation sold unproven stem-cell therapies to Italian people as a panacea for any number of conditions. A report into the affair published by the Italian Senate last week analyses what went wrong. It details the complicated history of the affair, and identifies the sorry list of characters — including many politicians — who must share the blame.

The report makes ten sensible proposals to fix the system. Politicians must now fine-tune and enact them. Stamina's is not the only unproven therapy to be indulged by the Italian state in recent times, but it needs to be the last. Scientists and policy-makers in other countries would be wise to take the lessons on board, too, because such lessons show that being right is not enough to stop harm from being done.

Most strikingly, the Senate investigation found no systematic failure in the state's technical agencies mandated to protect the public. The message from scientific experts in these agencies was loud and clear: that the Stamina claims had no merit and the technique carried considerable risks. But this expertise was ignored by other pillars of the state: the legislature and judiciary.

In 2012, the Italian Medicines Agency had declared the Stamina treatment unsafe, and ordered the closure of the company's laboratory in a Brescia hospital. But, egged on by Stamina's relentless campaigning, more than 450 people took to local courts and demanded the therapy anyway, on compassionate grounds. About half of them got it. Politicians, fearing an electoral backlash if they seemed to be ignoring patients, followed — even though they knew that a police investigation was under way.

The politicians went further. Despite formal advice from their own scientific agencies, they approved ministerial decrees to promote the Stamina treatment, including one that set up a state-sponsored clinical trial.

Scientists around the world looked on incredulously as the treatments continued until August 2014, when a Turin court finally ordered the confiscation of equipment and cells from the Stamina laboratory.

A small group of scientists from across Italy, including Cattaneo, fought relentlessly and at great personal cost against the politically powerful Stamina supporters. Their story is told in a 2014 *Nature* Comment (E. Cattaneo and G. Corbellini *Nature* **510**, 333–335; 2014). When Cattaneo was appointed to the Senate by former Italian president Giorgio Napolitano, this fight acquired another dimension. Her first action was to push for the Senate investigation. The 15-strong commission began work at the end of January last year, sifting through documents and holding 25 hearings to assess evidence that anti-Stamina campaigners would have been unable to find.

One of the report's proposals is that any future court that is asked to recommend an unauthorized treatment on compassionate grounds should have a representative of the health ministry and the state prosecutor there, to make sure that all arguments

“The Stamina case has been a disgrace to Italy, but it shows the influence that individual scientists can have in fighting anti-science forces.”

are heard. Another is to introduce ‘Daubert standards’ into Italian courts, to regulate the quality of scientific expertise. These standards, which are used in US courts, require judges to ensure that expert scientific testimony is based on knowledge that is the product of sound scientific methodology.

The report also proposes changing the decrees and regulations that relate to compassionate use of unauthorized therapies, to

close loopholes that could lead to abuse. It suggests new rules to ensure that ethics committees are truly independent, and recommends guidelines for media reporting similar to those adopted by the BBC last year.

Italian scientists are often disheartened by the lack of respect for science in their country. They are still reeling from the manslaughter verdict meted out by a court to seismologists who had advised the government before the major 2009 earthquake in L'Aquila; the convictions were overturned on appeal last year. Moreover, the Stamina case has a close parallel in the notorious case of Luigi Di Bella, a physician who claimed in the 1990s that a mixture of molecules such as somatostatin, together with vitamins, could cure cancer.

The Stamina case has been a disgrace to Italy, but it shows the influence that individual scientists can have in fighting — even against seemingly impossible odds — anti-science forces. And as if to underline the point that science can prevail in the most hostile of environments, a day after publication of the Senate report, the European Commission formally authorized the Western world's first-ever approval for a stem-cell therapy: a treatment for a rare type of blindness that has been developed entirely by Italian scientists, working exclusively in Italy.

It is not just in the political world that researchers can help others to see more clearly. ■

No strings

Details of a climate-change sceptic's links to the energy industry make worrying reading.

Earlier this week, documents were passed to the news media, including *Nature*, that link the global-warming sceptic Willie Soon, a researcher at the Harvard-Smithsonian Center for Astrophysics (CfA) in Cambridge, Massachusetts, to funders in the energy industry and conservative circles.

The files include research contracts and year-end reports, and provide new details on the kind of “deliverables” that Soon was providing for his funders. The group that released them, the Climate Investigations Center in Alexandria, Virginia, raised legitimate questions about whether Soon had properly disclosed this funding to journals that published his work. The CfA has responded by launching an investigation.

Willie Soon has been a poster child for the small community of climate-change sceptics for more than a decade. Environmentalists and other watchdogs have examined and exposed his industry funding countless times. Unknown until now were the explicit details of

the strings that come with such funding. In some cases, these strings included requirements that Soon show copies of proposed publications to Southern Company, a major electric utility that has given him nearly US\$410,000 since 2006, for input before publication. The company did not have the right to require changes, but another provision prevented Soon and the CfA from revealing its involvement without prior notification. This is troubling indeed.

Many scientists receive funding from industry as well as from foundations. Private money for science can often have an agenda, and this is why transparency matters so much. (Global-warming sceptics say that government funding has the same taint, but this comes with an assumption of disclosure.)

CfA director Charles Alcock said that agreeing to provisions to limit disclosure was a mistake, and one that the centre will not repeat. Although it has no explicit rules on disclosure, the centre does expect scientists to follow the publishing rules that journals set out.

One thing does not add up. The CfA, after all, is launching an investigation into one of its own staff members on the basis of the evidence of its own documents, but only after it was forced to hand them to an environmental group under a Freedom of Information Act request. Whether or not Soon fully disclosed the source of his funding to all of the journals remains unclear, but the basic facts were always there.

Alcock says that his job is to protect academic freedom at all costs. Fair enough. But freedom comes with responsibilities. ■

A sore thing

The use of technologies that objectively measure pain must be carefully monitored.

Injuries and illness evoke sympathy, so why do we find it so hard to appreciate another's pain? Why, as William Shakespeare observed, when we encounter: “A wretched soul, bruised with adversity” do “We bid be quiet when we hear it cry”? One answer is that there is no objective way to measure pain. This is especially true for the enduring nature of chronic pain, when the original physical injury — if there was one — is long gone.

For the millions of people worldwide who truly suffer from some type of pain, such scepticism means that they go untreated. The problem is particularly acute for women who, for reasons that are poorly understood, are far more likely to experience chronic pain than men. Yet the stereotype of the ‘hysterical’ woman persists. Women are substantially less likely than men to receive prescription opiates; instead they receive sedatives or antidepressants.

The signal of pain sits in the brain, and researchers are inching towards ‘pain-o-meter’ brain scans that could replace or complement a person's subjective self-reporting of their suffering. An objective measure of the brain activity that accompanies chronic pain could go a long way to changing the public perception that people without obvious physical injuries are imagining or faking their condition.

As we explore in a News Feature on page 474, an increasing number of lawyers want to see the technique introduced as evidence in court, to help injured clients to prove that they are not malingering. Start-up companies are charging ahead to offer commercial scans as documentation.

This development makes many scientists and ethicists nervous. By scientific standards, many of the methods have not yet been tested on enough people to prove that they are accurate and impossible to cheat. In response, lawyers argue — fairly enough — that even if the tests are not statistically indisputable, there is no harm in providing one more piece of evidence to back up their clients' claims.

Much more worrying is the possibility that the technology will be

misused, forcing plaintiffs — or even patients — to prove that they are in pain to receive compensation, insurance coverage or pain medication. Although it is unlikely that physicians will begin routinely ordering expensive scans of their patients' brains before prescribing opiates, it is easy to imagine insurance companies wanting proof of chronic pain before they shell out for years of treatment.

Pain-o-meters have a research use as well. Several major pharmaceutical companies are already beginning to use such neuroimaging techniques to test new painkillers — a notoriously difficult task because of the myriad threads that create the mental tangle of the pain experience. Fear, depression, attention and the power of suggestion all colour a person's report of sensation. The end result is that promising analgesic drugs are thrown out because patients might not think they are helping, even if they are actually fixing the biological cause of the pain. An objective measuring stick could allow researchers to push this jumble aside, treat the pain, and then treat the factor that is making the patient think that they are still in pain. At the same time, a better understanding of what pain looks like could reveal new drug targets.

Discussion of chronic pain mirrors other debates in medicine, most notably the distinction frequently drawn between physical and mental suffering. Legal systems and society as a whole persist with the idea that mental anguish is somehow different and less important. US courts, for instance, allow compensation to be paid for physical injury, but rarely emotional injury.

Laws and attitudes have simply not evolved with the scientific understanding of the brain. The idea that illnesses such as depression or post-traumatic stress disorder are the result of physically disordered brain circuits is catching on. Neuroscientists are comfortable blurring the line between the physical and the mental as they search for the biological roots of disease.

Pain does more than blur the lines between the mental and the physical; it unites them. Because each individual's experience is the product of so many components, brain scans may not pick up what feels very real to the sufferer. Scientists know surprisingly little about how exactly chronic pain intertwines with emotional and mental processes, which seem to be responsible for perpetuating the feeling long after the injured nerves have healed.

Measuring pain might not make it go away, but it could still offer some relief. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunqf



Focus on political Islamic groups to boost science

For science to realize its potential in the Muslim world, attitudes need to change at a societal level, not just an individual one, says Dyna Rochmyaningsih.

Recent terrorist attacks in Europe and the continued activity of the jihadist group ISIS in the Middle East have thrown the spotlight firmly back on radical Islam. Some studies blame the Muslim world's poor and unstable economies for the spread of this fundamentalism. Presumably then, improving the economy could help Muslim societies to tackle these radical movements.

Science can play a big part in this economic development, as it has in other places. But because some Muslims see a conflict between science and their faith, the philosophical question of how to reconcile the two is at the heart of many efforts to advance scientific development in the Muslim world.

Earlier this month, the organization Muslim-Science.com gathered a task force of prominent Muslim scholars in Istanbul to discuss the importance of reconciliation to "the future of the Islamic Project and its ability to embrace modernity".

The marriage of faith and science produced advances in mathematics, medicine, physics and astronomy during the Middle Ages. To recreate the enlightened attitudes present during this golden age, the scholars argue, Muslim scientists need to build broader societal support for science.

In my view, this focus on personal reconciliation is a naive way to address the problem and one that is unlikely to have much effect. Reconciliation is philosophically and theologically important for individual scientists, but will have little impact on wider society. It demands critical thinking. And, ultimately, the scientists it concerns form a tiny part — about 0.01% — of the world's Muslims.

Attempts at reconciliation could even make the situation worse, and harden anti-science attitudes in Muslims. That has happened in Indonesia, home to the world's largest Muslim population, with the publication of the book *Adam Was Born*, which attempts to reconcile Islamic faith and evolution. By re-analysing verses of the Koran as positive to science, the author, Agus Mustofa, enraged traditional clerics and polarized opinions.

The problem is that, unlike Catholicism, Islam has no unifying voice of authority to rule on koranic interpretation. Although it is legal for Mustofa to interpret the Koran, clerics have much more influence, and this gives them great power.

Rather than reconciliation, it is important to monitor and understand the way in which political and ideological groups influence how young Muslims view science.

The radical Islamists of ISIS see science as an attribute of their enemies. They have denounced the great Medieval Muslim scientists Ibn Sina and Ibn al-Nafis as heretics and atheists. It is clear that such rhetoric — if influential — will hold back scientific development in Muslim countries.

Here in Indonesia, for example, groups such as the Muslim Brotherhood and Hizbut Tahrir have a strong presence in high schools and universities, and this gives them profound influence on young Muslims' views of the world, including science.

The influence is not all negative to science. The Muslim Brotherhood, although hostile to evolution, encourages talented scientists to develop their careers and helps to place them on postgraduate courses overseas, typically in Japan. Many of these people return to Indonesia as university lecturers.

However, some Muslim groups think that asking a lot of questions is a Jewish trait, and one not to encourage. The convener of the task force, Usama Hasan, says that just as the Nazis labelled quantum mechanics as Jewish science, so fundamentalist Muslim groups talk about *kafir* science — the science of the unbeliever.

These groups have much more potential to influence the future scientists, engineers and politicians of the Muslim world than individual researchers. Yet Muslim scholars have largely ignored them.

The organizers of the Istanbul event, for example, also held a meeting on science education in the Muslim world. Scholars at the meeting have proposed recommendations that will be submitted in June to the Organisation of Islamic Cooperation. None mentions the potential of political and ideological groups.

The *Atlas of Islamic-World Science and Innovation*, a report initiated by Britain's Royal Society and published last December, says that science and technology is being held back by the same issues in the Muslim world as in many developing

nations: poor funding, low investment in people and a lack of international collaboration.

This report, too, ignores the potential of political Islamic groups. As a result, it makes the same recommendations for improving science in the Muslim world as for, say, South America. We must look at what makes each region unique.

There is no easy way to counter the impact of political Islamic groups on science, but it should be studied and accounted for. And it should certainly take priority over the reconciliation of science and faith.

Reconciliation is an individual process, and something that is intangible in the realm of policy-making. By contrast, hard-line groups can influence whole societies. To capitalize on this influence, we might need to reform science education in primary schools in the Muslim world, and teach young people to think for themselves before they are exposed to political ideas. ■

RECONCILIATION
IS AN
INDIVIDUAL
PROCESS, AND
SOMETHING THAT IS
INTANGIBLE
IN THE REALM OF
POLICY-MAKING.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/bd9ngj

Dyna Rochmyaningsih is a freelance science journalist in Jakarta.
e-mail: drochmya87@gmail.com

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

PHOTONICS

Water lens with adjustable focus

Researchers have developed a microscopic lens with a focal length that can be controlled in less than a millisecond.

Controlling the focus of an optical lens is useful for microscopy and photography, but existing reconfigurable lenses are often bulky or slow to adjust. Romain Quidant and his colleagues at the Institute of Photonic Sciences in Barcelona, Spain, created a controllable lens by placing a disc of gold nanorods inside a thin chamber of water and putting it on top of a conventional lens.

They used a laser to excite the electrons in the nanorods, heating the water and changing its refractive index to create a lens-like effect. The team was able to vary the focal distance of the lens by tens of micrometres with sub-nanometre accuracy, and in only 200 microseconds. **ACS Photonics** <http://doi.org/2cd> (2015)

BIOMATERIALS

DNA-based gel for printing organs

A gel that can be infused with live cells and nutrients makes a promising material for printing three-dimensional tissues such as artificial organs.

Dongsheng Liu at Tsinghua University in Beijing, Wenmiao Shu at Heriot-Watt University in Edinburgh, UK, and their team made two water-based inks from peptides and synthetic DNA strands



NOBUO IWATA/MOMENT OPEN/GETTY

ECOLOGY

Competing bluebirds make tougher sons

Female western bluebirds that have to compete for nesting sites produce more early-hatching male chicks than do females with fewer competitors. The chicks are also likely to be more aggressive. This has long-term effects on the range and behaviour of subsequent generations.

Renée Duckworth and her colleagues at the University of Arizona in Tucson discovered that female western bluebirds (*Sialia mexicana*; pictured) that live in areas with many neighbours and few nesting sites laid eggs containing more androgen — a hormone that boosts aggression

in the offspring — than females facing less competitive pressure. Those first eggs also tended to produce more males, which can compete for and colonize new territory. When the researchers increased the number of nesting sites in study areas in western Montana, however, the females produced eggs with less androgen, and fewer male offspring in the early eggs.

This eventually allowed the western bluebird to boost its numbers and displace its competitor, the mountain bluebird (*S. currucoides*).

Science 347, 875–877 (2015)

VOLCANOLOGY

Sulfur in magma gets a lift

Sulfur and metals can hitch a ride on bubbles rising in molten magma. This could explain why some volcanoes spew out more sulfur than expected, and how metal ores can form in the crust nearby.

Sulfur-rich magma normally sinks to the bottom of magma chambers. A team led by Jim Mungall at the University of Toronto in Canada used lab studies and

mathematical modelling to show that magma droplets, which contain metals, can form on the surface of vapour bubbles. Droplets that do not reach the surface cool and form rocks that are rich in sulfur, copper and gold.

In another study, Jon Blundy and his team at the University of Bristol, UK, used lab experiments to conclude that sulfur-rich gases interact with salty, copper-rich fluids inside a magma chamber to form thick deposits of copper-based minerals — similar to those that provide three-quarters of

DONGSHENG LIU AND WENMIAO SHU, ANGEW. CHEM. INT. EDN

the world's copper.
Nature Geosci. <http://dx.doi.org/10.1038/ngeo2373>;
<http://dx.doi.org/10.1038/ngeo2351> (2015)

EPIDEMIOLOGY

Plague came to Europe in waves

The bacterium that causes the plague, which killed millions of Europeans over four centuries from the 1350s, was repeatedly reintroduced from Asia and did not establish itself in European rodents as was thought.

Yersinia pestis bacteria live in wild rodents and can infect humans when climate changes cause rodent populations to collapse, triggering plague-carrying fleas to find alternative hosts. To locate plague reservoirs in Europe, Nils Christian Stenseth at the University of Oslo and his colleagues analysed historical outbreaks along with tree-ring-based records of climate. They found no connection between fluctuations in European climate and plague outbreaks, but did find links between Asian climate changes and outbreaks at European trade harbours.

The authors conclude that the plague took about 15 years to travel overland to Europe.
Proc. Natl Acad. Sci. USA
<http://dx.doi.org/10.1073/pnas.1412887112> (2015)

BIOCHEMISTRY

Sunlight damages DNA in the dark

Sunlight can cause cancer-related DNA damage hours after light exposure, owing to a skin pigment that was largely thought to be protective.

Douglas Brash at Yale University School of Medicine in New Haven, Connecticut, and his team studied how the pigment melanin in mouse skin cells responds to ultraviolet (UV) light. They found that UVA radiation, the main type of UV light that comes from the Sun and from tanning beds,

creates melanin by-products that damage DNA, generating DNA derivatives called cyclobutane pyrimidine dimers (CPDs) for up to three hours after light exposure.

CPDs are associated with the skin cancer melanoma, so blocking their formation could be a way to develop sunscreens that can be used after exposure to sunlight, the team says.
Science 347, 842–847 (2015)

CANCER

Bacteria protect tumours

Bacteria hiding out in tumours can shield them from attack by the immune system.

The oral bacterium *Fusobacterium nucleatum* has been linked to premature birth, rheumatoid arthritis and colon cancer. Gilad Bachrach and Ofer Mandelboim at the Hebrew University of Jerusalem and their colleagues studied the impact of the bacterium on cancer cells. They found that *F. nucleatum* sticks to tumour cells grown in culture and inhibits immune cells by activating an immune-cell receptor called TIGIT. Many immune-cell types found in human colon cancer and melanoma samples also expressed TIGIT, and were inhibited by *F. nucleatum*.

The results could explain why certain tumours, especially intestinal ones, seem to have high levels of bacteria.
Immunity 42, 344–355 (2015)

PLANT SCIENCE

Nectar fends off bee parasites

Floral nectar helps to control parasites in bumblebees.

Plants produce molecules called secondary metabolites that are harmful to herbivores but in some cases can also protect animals from parasites. To see whether such metabolites in nectar similarly affect pollinators, Leif Richardson at Dartmouth College in Hanover, New Hampshire, and his team

SOCIAL SELECTION

Popular articles on social media

Scientists cautious about outreach

Scientists think that they should actively participate in public debates about science and technology — but many have misgivings about doing so, according to a survey of nearly 4,000 US researchers. The results of the poll, by the Pew Research Center, inspired a fresh online conversation about the use of social media in public engagement. “Been saying for years scientists need to come down from ‘ivory tower’ and engage public,” tweeted Caleph Wilson, a cancer researcher at the University of Pennsylvania in Philadelphia. Ajinkya Kamat, a physics PhD student at the University of Virginia in Charlottesville, tweeted: “We need more avenues, better incentive structure to get scientists in all career stages involved in science outreach.”

➔ **NATURE.COM**
For more on popular papers:
go.nature.com/8mfcqm



infected eastern bumblebees (*Bombus impatiens*) with an intestinal parasite and gave the bees one of eight different nectar compounds. Four of the metabolites reduced the load of parasites by 60–80%.

The compound with the strongest effect on parasites, anabasin, did not seem to boost bumblebee survival, but the team says that these chemicals in nectar could benefit the bee colony as a whole by reducing parasite spread.

Proc. R. Soc. B 282, 20142471 (2015)

PALAEOECOLOGY

Coral growth shut down for millennia

Coral reefs in the eastern Pacific Ocean stopped growing for 2,500 years, probably because of a change in climate four millennia ago.

Lauren Toth at the Florida Institute of Technology

in Melbourne and her colleagues extracted a 2.68-metre core from a reef in the Gulf of Panama (pictured), representing 6,750 years of growth. They analysed the chemical composition of 133 skeletons of *Pocillopora* corals in the sample to assess coral health, local temperature, ocean currents and rainfall. They found that roughly 4,100 years ago, cooler temperatures and greater rainfall — similar to today's La Niña weather systems — were associated with the beginning of a 2,500-year pause in coral growth. The health of the corals seems to have declined at the start of this hiatus.

The samples also suggest that temperature is a key factor affecting coral growth.

Nature Clim. Change <http://dx.doi.org/10.1038/nclimate2541> (2015)

➔ **NATURE.COM**
For the latest research published by Nature visit:
www.nature.com/latestresearch

SEVEN DAYS

The news in brief

PEOPLE

IPCC head resigns

Rajendra Pachauri resigned as head of the Intergovernmental Panel on Climate Change (IPCC) on 24 February, amid allegations of sexual harassment. The accusations were made by a colleague at The Energy and Resources Institute (TERI), the non-profit organization that Pachauri directs in New Delhi, India. The IPCC said that panel vice-president Ismail El Gizouli would step in as acting chair for its session in Nairobi this week. Pachauri became chair of the IPCC in 2002, and was scheduled to complete his second term in office in October. See go.nature.com/1ssogm for more.

US data chief

The White House appointed its first-ever chief data scientist on 18 February. DJ Patil, a former mathematician who helped to coin the term 'data science' and who has worked at companies including Skype, PayPal and eBay, will be in charge of US government policies around open data, based at the Office of Science and Technology Policy. He will also work on the US Precision Medicine Initiative, announced in January, which seeks to link genomic data to health records in order to find patient-customized treatments.

Financial conflict

The Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, has launched an investigation into solar physicist and climate-change sceptic Willie Soon, after documents detailing research contracts with the energy industry and a conservative foundation were released in response to a US Freedom of Information



KATE HOLT/EVINE

Nations asked to tackle tropical diseases

Neglected tropical diseases affect more than 1.5 billion people worldwide, yet many diseases, such as river blindness may be prevented simply by taking a pill. In a report released on 19 February, the World Health Organization

estimated that it will cost US\$34 billion over 16 years to meet its targets to reduce the burden of the 17 neglected tropical diseases, which include leishmaniasis and leprosy. It has called on affected countries to boost their spending.

Act by Greenpeace. Officials at the Climate Investigations Center in Alexandria, Virginia, which revealed the documents on 21 February, allege that Soon failed to report these financial relationships on numerous peer-reviewed papers. See go.nature.com/tvgedg for more.

X-ray pioneer dies

Ernest Sternglass, a pioneer of safer X-ray imaging, died on 12 February, aged 91. Born in Berlin in 1923, Sternglass fled Germany to the United States in 1938. While at the firm Westinghouse, he worked on electron-amplification effects that were later harnessed in

the low-light television camera that allowed viewers to watch the live Moon landing in 1969. In the 1980s, while at the University of Pittsburgh, Pennsylvania, Sternglass and colleagues developed digital X-ray imaging. He also applied his studies on the dangers of X-ray exposure to the health effects of atomic-bomb testing, about which he publicly campaigned.

POLICY

Future of graphene

Europe's €1-billion (US\$1.3-billion) initiative to commercialize graphene, which has 142 industrial and academic partners in

23 countries, is on course and "providing excellent value for money", its organizers said on 24 February. In mid-January, an independent assessment of milestones reached by the Graphene Flagship project, produced for the European Commission, gave positive scores all round, a spokesperson for the project said. That assessment has not yet been published, but a 200-page road map setting out research areas for graphene and other two-dimensional crystals was published this week, covering 11 themes from energy storage to biomedical devices (A. C. Ferrari *et al.* *Nanoscale* <http://doi.org/2df>; 2015).

POLICY

Dietary advice

The US Dietary Guidelines Advisory Committee released its scientific assessment of nutritional guidance on 19 February. The report calls for the eradication of limits on dietary cholesterol, sets an upper limit for added sugar consumption, and recommends a diet containing more vegetables and less meat than are consumed by many people — both to improve health and to reduce environmental impact. The scientific assessment is available for public comment until 8 April. The final publication, due out in the autumn, will be used to guide health recommendations as well as public programmes such as school lunches and food assistance.

EVENTS

Indian PhDs protest

Thousands of Indian PhD students are protesting (pictured) about delays in a hike in research-fellowship wages pledged by the Indian government last October, after protests in July. Last week, half a dozen students went on hunger strike. While some funding agencies have implemented the rise, others have delayed its introduction, prompting a letter from students to Prime



Minister Narendra Modi in January. See go.nature.com/adq5pw for more.

New killer virus

The US Centers for Disease Control and Prevention announced the discovery of a new deadly virus on 19 February. Dubbed Bourbon virus after the county in Kansas where it was found, it is thought to have killed a man, aged over 50, who was bitten by ticks shortly before falling ill. Bourbon virus is part of a family known as thogotoviruses, which are carried by ticks and insects and are known to have infected only eight people previously.

Anti-HIV-jab trial

Trials to test whether injectable antiretroviral drugs could prevent HIV infection for months at a time were launched

by the US National Institutes of Health on 19 February. The trials, to be carried out in Africa, South America and the United States, will compare the effectiveness of the injections against placebo and oral forms of the medications. Currently, one pill (Truvada) is approved for HIV prevention in the United States, but it must be taken daily; injectable drugs might offer long-lasting protection.

Fast Ebola test

The first rapid diagnostic test for Ebola — which gives a result in 15 minutes — was approved by the World Health Organization in Geneva, Switzerland, on 20 February. The inexpensive test, developed by Corgenix of Broomfield, Colorado, and Robert Garry, a virologist at Tulane University in

New Orleans, Louisiana, detects viral protein and does not require electricity or refrigeration, making it suitable for use in remote settings. Currently, blood samples from suspected cases in West Africa must be transported to labs for genetic testing, which incurs delays. Faster diagnosis is considered key to combating the epidemic. See go.nature.com/e5bml6 for more.

FUNDING

Costly drugs

The National Health Service (NHS) in England is doing more harm than good by paying for expensive new drugs, says a team of UK health economists. The researchers suggest that the threshold used by the National Institute for Health and Care Excellence to assess the cost effectiveness of new drugs should be drastically lowered (K. Claxton *et al. Health Technol. Assess.* **19**, 14; 2015). The NHS will currently fund treatments that cost up to £30,000 (US\$46 000) for every extra year of good-quality life they provide. If this figure was dropped to £13,000, benefits would spread to more people, say the authors.

Plutonium shortage

It is likely to take longer than anticipated to bolster NASA's dwindling stocks of plutonium-238, which is used to power deep-space missions. The US Department of Energy (DOE) has been ramping up production of the radioactive isotope, aiming to provide NASA with more than 1 kilogram per year by 2021 (see *Nature* **515**, 484–486; 2014). But limited funding means that it will take an unspecified time longer. DOE space-power director Alice Caponiti told a NASA advisory panel on 20 February. With limited supplies, NASA has had to ration the fuel.

NATURE.COM
For daily news updates see:
www.nature.com/news

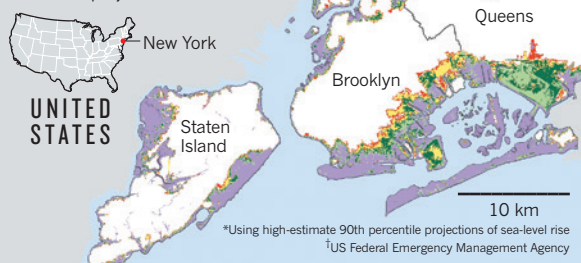
TREND WATCH

Global warming is likely to increase sea levels around New York City by 56–127 centimetres by 2100, and in the worst case by 183 cm, warns an assessment released by the New York City Panel on Climate Change on 17 February. The study says that the city's sea levels have risen by around 3 cm per decade — nearly twice the global average — since 1900. The coastal area affected by floods could double by 2100, with the boroughs of Queens and Brooklyn having the largest amount of land area at risk.

FUTURE FLOOD ZONES

Sea-level rise in the coming century is projected to put the people, economy and infrastructure of New York City at great risk.

Predicted sea-level rise*:
2020s: 25 cm 2080s: 147 cm
2050s: 76 cm 2100s: 191 cm
December 2013
FEMA[†] projections



NEWS IN FOCUS

SCIENCE IN THE CITY Francis Crick Institute worried about railway **p.464**

ARTIFICIAL INTELLIGENCE Game-playing 'bot' is test bed for human intelligence **p.465**

DWARF PLANETS Graphical guide to Pluto and Ceres missions **p.468**



NEUROSCIENCE Brain imaging for pain heads for the courtroom **p.474**

JANINE WIEDEL PHOTOLIBRARY/ALAMY



of associations — for example, deducing that smoking during pregnancy is linked to poor child development, and that children born at socio-economic disadvantage are more likely to struggle at school.

Researchers argue that new birth cohorts are needed. Children born today, at least in most Western countries, enter a world that is increasingly warmer, more digitized, more ethnically diverse and more obese, with wider income inequality, than it was even a decade ago. New questions and techniques, such as sophisticated genetic analyses, also arise as time goes on, allowing different information to be gleaned.

The National Children's Study aimed to follow 100,000 children from birth to age 21, but was cancelled in December 2014 before it fully launched, 15 years and US\$1.2 billion after its inception (see *Nature* <http://doi.org/2dh;2014>). Scientists had started to recruit parents and children, but the study struggled to find a clear scientific direction, had trouble enrolling participants and racked up eye-watering costs.

Meanwhile, scientists in the United Kingdom have been getting their own birth-cohort study off the ground, although it has attracted much less attention than the US study. The team involved, led by paediatric epidemiologist Carol Dezateaux of University College London's Institute of Child Health, officially launched the Life Study this week at the House of Lords, to raise its profile among politicians and policy-makers.

Studies in Norway and Denmark are each currently following more than 100,000 children, and the United Kingdom already has a series of smaller birth cohorts, the first of which started in 1946 (see *Nature* **471**, 20–24; 2011). But the Life Study aims to distinguish itself, in particular by collecting detailed information on pregnancy and the first year of the children's lives — a period that is considered crucial in shaping later development.

The scientists plan to squirrel away freezer-fulls of tissue samples, including urine, blood, faeces and pieces of placenta, as well as reams of data, ranging from parents' income to records of their mobile-phone use and videos of the babies interacting with their parents.

The idea of a major new British birth cohort was first aired in the mid-2000s, but it took years to get organized. Government funding bodies agreed in 2011 to pay £38.4 million (US\$60 million) until 2019. The scientists have since done pilot studies, and late last autumn they started to recruit parents into the study proper, aiming ▶

Birth-cohort studies follow children from the moment they are born to learn what shapes their life trajectory.

EPIDEMIOLOGY

Wanted: 80,000 British babies

UK launches massive study to track children from birth, months after closure of US counterpart.

BY HELEN PEARSON

An ambitious study that will follow 80,000 children from cradle to grave has launched in the United Kingdom, two months after a similar project in the United States ended in expensive failure.

The project aims to track a generation of

twenty-first-century babies and work out which factors in their early lives are important in shaping their health and wealth as they grow into adults. There are reasons to hope that the Life Study will have a happier ending than its US counterpart, the National Children's Study.

Such 'birth-cohort' studies are prized. Scientists have used them to extract a stream

► to enrol all participants by 2018.

Certain factors make researchers optimistic that the British study will succeed where the US one failed. One is the National Health Service, which provides care for almost all pregnant women and their children in the United Kingdom, and so offers a centralized means of recruiting, tracing and collecting medical information on study participants.

In the United States, by contrast, medical care is provided by a patchwork of different providers. “I think that most researchers in the US recognize that our way of doing population-based research here is simply different from the way things can be done in the UK and in Europe, and it will almost always be more expensive here,” says Mark Klebanoff, a paediatric epidemiologist at Nationwide Children’s Hospital in Columbus, Ohio, who was involved in early discussions about the US study.

At one stage, US researchers had planned to knock on doors of random houses looking for women to enrol before they were even pregnant. “It became obvious that that wasn’t going to be a winning formula,” says Philip Pizzo, a paediatrician at Stanford University in Palo Alto, California, who co-chaired the working group that concluded that the National Children’s Study was not feasible. “The very notion that someone was going to show up on your doorstep as a representative from a government-funded study and say ‘Are you thinking of getting pregnant?’ was not so attractive sociologically.”

Researchers involved in the UK study say that they hope to learn from the challenges faced by their US counterparts — they have a clear study design and recruitment strategy — and that they are keen to collaborate internationally. The major concern is whether enough interested parents will sign up, something that will become apparent only in the next few months. “It’s the known unknown,” says Dezateux.

US researchers mourn the demise of their study. “We have now lost the opportunity to remain at the forefront of this field, and to collect the crucial life-course data,” says Ezra Susser, an epidemiologist at Columbia University in New York City. But starting a study that lasts a lifetime comes with particular challenges wherever it is done, says Pizzo. For instance, about one-third of the children in the UK study are expected to live to 100. The scientists designing the study will be long dead by then, and can only hope that the information they collect will still be useful. “The responsibility of getting it right will be enormously significant,” says Pizzo. “If you think of what’s happened in the last decade — in terms of social media, how we connect, the insights around basic biology — 100 years from now, it’s almost imponderable to think where knowledge is going to be.” ■



The Francis Crick Institute sits at the nexus of three central London railway hubs.

URBAN SCIENCE

Biology powerhouse raises railway alarm

Central London’s Francis Crick Institute fears that proposed train line will disrupt delicate science experiments.

BY DANIEL CRESSEY

A landmark addition to London’s science scene is on a collision course with the expansion of the city’s transport system. The Francis Crick Institute warns that vibrations and electromagnetic fields generated by Crossrail 2, a proposed railway line that would skirt the institute, could interfere with scientific work there. The teams behind both efforts are now seeking a solution.

Set to employ 1,250 scientists and to have an annual budget of more than £100 million (US\$154 million), the Crick, as it is known, is destined to become one of Europe’s medical-science powerhouses. Construction is due to be completed in November. The warnings about Crossrail 2 first emerged from the UK Medical Research Council (MRC), which provided the lion’s share of the Crick’s construction budget and will move staff from its National Institute of Medical Research (NIMR), in another part of London, to the new facility.

In public documents submitted to a UK parliamentary committee and discussed at the MRC over the past year, the council warned of “potentially serious consequences for operation of sensitive scientific equipment” if Crossrail 2 goes ahead, and said that “expensive remediation works” might be required.

The Crick has now sounded its own warning in a statement to *Nature*: “The Crossrail 2 trains stopping and starting at the proposed station would have an electromagnetic impact on our imaging facilities.” The imaging equipment — which includes nuclear magnetic resonance spectrometers, electron microscopes and super-resolution microscopes — will support “a significant component” of research at the building, says the Crick.

When the institute was first proposed, some scientists at the NIMR raised concerns about interference from multiple railway lines that already surround the site (see ‘Science in the city’). MRC chief executive John Savill says that the Crick’s design “fully took into account” vibration, noise and electromagnetic interference from these; the Crick team says that it deliberately located its imaging facilities away from the major Thameslink line. But in 2013, Crossrail 2’s proposed route was modified to allow the line to connect with Euston station, which, says the Crick, would place the railway too close to the imaging facilities.

The government will decide next month whether to use the proposed route in the next stage of planning; the Crossrail 2 team hopes to start construction in 2020 at the earliest.

“The Department for Transport has given assurances that the route selected will

LAING O’ROURKE/FRANCIS CRICK INST.

not impede any research activities due to interference from railway operations or construction works,” says Savill. Transport for London (TfL), the local government organization that is driving the Crossrail 2 plans, says that it takes the concerns “very seriously”. The MRC is now working with TfL to find a solution.

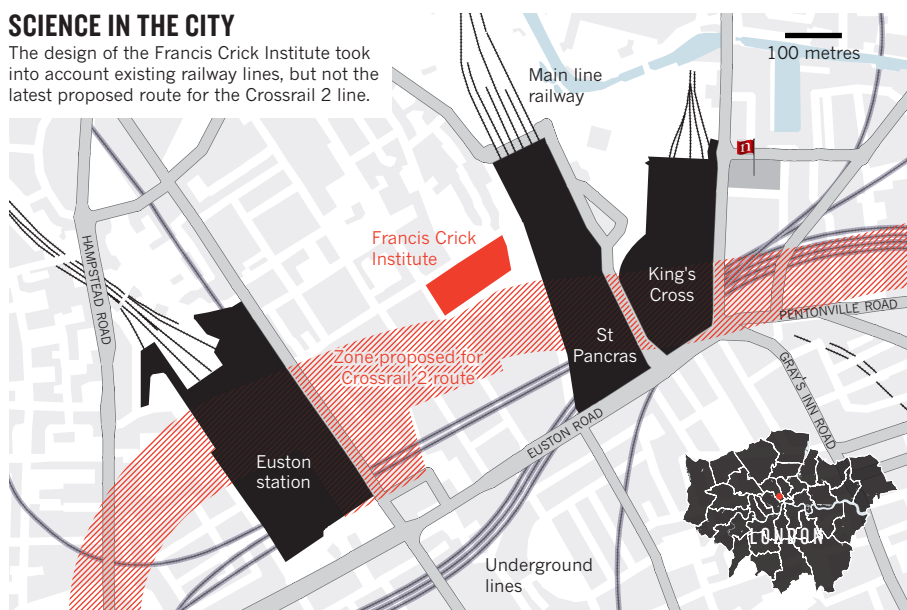
The Crick team would like Crossrail 2 to be diverted back to its former route, away from the institute. But Michèle Dix, managing director for Crossrail 2, says that plans have already changed because of concerns from the Crick. Among other things, the planned tunnel has been shifted farther underground. “You can’t just keep on moving it deeper and deeper,” she says, adding that further concerns should be addressed through engineering — for example by making the tunnel linings thicker.

Daniel Moylan, a TfL board member and a transport adviser to London mayor Boris Johnson, says that the mayor is a huge supporter of both Crossrail 2 and the Crick. The latter is a major element in Johnson’s plans to promote the capital as ‘MedCity’ — a global hub for life-sciences research. Moylan is confident that the institute’s “legitimate” concerns can be allayed by means of technical solutions.

Objections from Parisian academics in the first half of the twentieth century are said to have affected the route of the Métro, but there have also been more-modern conflicts between scientific equipment and transport infrastructure.

SCIENCE IN THE CITY

The design of the Francis Crick Institute took into account existing railway lines, but not the latest proposed route for the Crossrail 2 line.



In 2013, the proposed route of a US light-rail line was diverted after concerns from the University of Colorado Denver about the effects on spectroscopy and microscopes at its medical campus in Aurora. In 2011, the University of Maryland in College Park dropped its opposition to a new ‘Purple Line’ link to the Washington DC Metro once the local transportation agency agreed to bury and shield power lines.

Other labs have depended on engineering solutions. The New York Structural Biology Center, located near a number of subway lines, placed its most sensitive equipment on concrete slabs attached directly to Manhattan bedrock. It experiences no problems from vibration now, says executive director Willa Appel. If there is no solution in London for the Crick, Appel says, “tell them they’re welcome to come here”. ■

ARTIFICIAL INTELLIGENCE

DeepMind algorithm beats people at classic video games

Computer that learns from experience provides a way to investigate human intelligence.

BY ELIZABETH GIBNEY

DeepMind, the Google-owned artificial-intelligence company, has revealed how it created a single computer algorithm that can learn how to play 49 different arcade games, including the 1970s classics *Pong* and *Space Invaders*. In more than half of those games, the computer became skilled enough to beat a professional human player.

The algorithm — which has generated a buzz since publication of a preliminary version in 2013 (V. Mnih *et al.* Preprint at <http://arxiv.org/abs/1312.5602>; 2013) — is the first artificial-intelligence (AI) system that can learn a variety of tasks from scratch given only the same, minimal starting information. “The fact that you have one system that can learn several games,

without any tweaking from game to game, is surprising and pretty impressive,” says Nathan Sprague, a machine-learning scientist at James Madison University in Harrisonburg, Virginia.

DeepMind, which is based in London, says that the brain-inspired system could also provide insights into human intelligence. “Neuroscientists are studying intelligence and decision-making, and here’s a very clean test bed for those ideas,” says Demis Hassabis, co-founder of DeepMind. He and his colleagues describe the gaming algorithm in a paper published this week (V. Mnih *et al.* *Nature* **518**, 529–533 (2015); see also News & Views on page 486).

➔ **NATURE.COM**
For a video that
peeks inside the
offices of DeepMind:
go.nature.com/2kqata

Games are to AI researchers what fruit

flies are to biology — a stripped-back system in which to test theories, says Richard Sutton, a computer scientist who studies reinforcement learning at the University of Alberta in Edmonton, Canada. “Understanding the mind is an incredibly difficult problem, but games allow you to break it down into parts that you can study,” he says. But so far, most human-beating computers — such as IBM’s Deep Blue, which beat chess world champion Garry Kasparov in 1997, and the recently unveiled algorithm that plays Texas Hold ‘Em poker essentially perfectly (see *Nature* <http://doi.org/2dw>; 2015) — excel at only one game.

DeepMind’s versatility comes from joining two types of machine learning — an achievement that Sutton calls “a big deal”. The first, called deep learning, uses a brain-inspired

architecture in which connections between layers of simulated neurons are strengthened on the basis of experience. Deep-learning systems can then draw complex information from reams of unstructured data (see *Nature* 505, 146–148; 2014). Google, of Mountain View, California, uses such algorithms to automatically classify photographs and aims to use them for machine translation.

The second is reinforcement learning, a decision-making system inspired by the neurotransmitter dopamine reward system in the animal brain. Using only the screen's pixels and game score as input, the algorithm learned by trial and error which actions — such as go left, go right or fire — to take at any given time to bring the greatest rewards. After spending several hours on each game, it mastered a range of arcade classics, including car racing, boxing and *Space Invaders*.

Companies such as Google have an immediate business interest in improving AI, says Sutton. Applications could include how to best place advertisements online or how to prioritize

stories in news aggregators, he says. Sprague, meanwhile, suggests that the technique could enable robots to solve problems by interacting with their environments.

But a major driver is science itself, says Hasabis, because building smarter systems means gaining a greater understanding of intelligence. Many in computational neuroscience agree. Sprague, who has created his own version of DeepMind's algorithm, explains that whereas AI is largely irrelevant to neuroscience at the level of anatomical connections among neurons, it can bring insight at the higher level of computational principles.

Computer scientist Ilya Kuzovkin at the University of Tartu in Estonia, who is part of a team that has been reverse-engineering DeepMind's code since 2013, says: "The tricks we use for training a system are not biologically realistic. But comparing the two might lead to new ideas

"The tricks we use for training a system might lead to new ideas about the brain."

about the brain." A particular boost is likely to come from the DeepMind team's choice to publish its code alongside its research, Kuzovkin says, because his lab and others can now build on top of the result. "It also shows that industry-financed research goes the right way: they share with academia," he adds.

DeepMind was bought by Google in 2014 for a reported £400 million (US\$617 million), and has been poaching leading computer scientists and neuroscientists from academia, growing from 80 to 140 researchers so far.

Its next steps are again likely to be influenced by neuroscience. One project could be building a memory into its algorithm, allowing the system to transfer its learning to new tasks. Unlike humans, when the current system masters one game, it is no better at tackling the next.

Another challenge is to mimic the brain's way of breaking problems down into smaller tasks. Currently, DeepMind's system struggles to link actions with distant consequences — a limitation that, for example, prevented it from mastering maze games such as *Ms. Pac-Man*. ■

NEUROSCIENCE

Researchers seek definition of head-trauma disorder

Guidelines should assist in diagnosis of brain disease seen in retired American footballers.

BY HELEN SHEN

Dave Duerson suspected that something was wrong with his brain. By 2011, 18 years after the former American football player had retired from the Phoenix Cardinals, he experienced frequent headaches, memory problems and an increasingly short temper. Before he killed himself, he asked that his brain be donated for study.

Researchers who examined it found signs of chronic traumatic encephalopathy (CTE), a degenerative condition linked to repeated head injuries. At least 69 cases have been

reported in the literature since 2000, many in former boxers and American football players (P.H. Montenegro *et al.* *Alz. Res. Ther.* 6, 68; 2014) — heightening public concern about concussions during contact sports. Yet much about CTE is unknown, from its frequency to its precise risk factors and even whether its pathology is unique.

Researchers now hope to take a major step towards answering those questions. At Boston University in Massachusetts on 25–27 February, neuroscientists will convene to examine the characteristics of CTE in brain tissue from post-mortem examinations. They hope to agree

on a set of diagnostic criteria for the disease, and to assess whether it is distinct from other brain disorders, such as Alzheimer's disease.

The effort is sorely needed, says Walter Koroshetz, acting director of the US National Institute of Neurological Disorders and Stroke in Bethesda, Maryland, which is organizing the meeting. "The definition is the important piece that lets you do the rest of the research," he says. And the stakes are high. CTE is associated with memory loss, irritability, depression and explosive anger, which are thought to appear and worsen years after repeated head trauma. Research by Ann McKee, a neuropathologist



**MORE
ONLINE**

HISTORY IN PICTURES

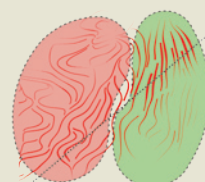


A celebration of the 85th anniversary of Pluto's discovery go.nature.com/vjl2ci

MORE NEWS

- Low leaks from US gas fields go.nature.com/8cmvu6
- FDA ponders rules on DNA testing go.nature.com/3ncdbh
- Star skirted Solar System 70,000 years ago go.nature.com/d6kxb6

VIDEO OF THE WEEK



How countries 'flow' towards development go.nature.com/8hjlsu



Repeated head injuries in American football have been linked to a degenerative brain disorder later in life.

JAMIE SQUIRE/GETTY

at Boston University, suggests that CTE often occurs alongside Alzheimer's disease and other neurodegenerative disorders. "I think this disease is more common than traditionally recognized," she says.

Such research helped to spur a major lawsuit against the US National Football League (NFL) over its handling of head trauma, involving some 4,500 former players and a potential settlement of more than US\$765 million. The findings have led researchers worldwide to examine the risks of head trauma in other sports, including professional rugby and amateur soccer.

To make CTE studies more systematic, McKee has proposed pathological criteria and disease stages defined by the severity of a person's symptoms (A. C. McKee *et al.* *Brain* <http://doi.org/2dc;2012>). CTE often involves deterioration of white matter in the brain, as well as abnormal aggregation of a DNA-binding protein called TDP-43 — which is also implicated in amyotrophic lateral sclerosis. But McKee says that CTE's most striking feature is the clumped distribution of an abnormal, hyperphosphorylated form of a protein called tau that normally helps to stabilize the internal structure of cells.

Tau deposits are found in the brains of elderly people and in those with Alzheimer's disease — leading some sceptics to question whether CTE is a distinct disorder. And critics such as Christopher Randolph, a neuropsychologist at Loyola University Medical Center in Maywood, Illinois, worry that CTE research has a basic sampling bias. The brain banks that support many CTE studies largely rely on tissue donated by players or their families, who often suspect that the individual had a severe neurological problem. This tips the balance towards those who may have various brain disorders.

Mistaking other conditions for CTE could have serious consequences, says Randolph, who worked with the Chicago Bears NFL team until 2002. "If you are in the grips of major depression, your actions may be different if you believe

you have a fatal neurodegenerative disease and you're just going to get worse, versus believing that you have a treatable illness," he says.

McKee acknowledges the sampling bias but says that acquiring donated tissue from healthy individuals will take time, because relatively few people request post-mortems in the absence of a visible problem. But she is convinced that many of the athletes' brains that she has received show an unmistakable signature that includes concentrations of hyperphosphorylated tau around blood vessels, especially in the furrows of the cerebral cortex. This is a departure, she says, from the more-uniform tau deposition seen in diseases such as Alzheimer's.

At the meeting, researchers will put these potential CTE signatures to the test. McKee and her team have prepared images of tissue slices from 25 brains from multiple brain banks. The sample, from players and non-players, includes suspected cases of CTE and known cases of Alzheimer's disease and other tau-based disorders. Neuropathologists will evaluate the images without information on the subjects' personal or medical histories, and will present their blinded diagnoses at the event.

If reliable neuropathological criteria can be found for CTE, researchers hope to use them to re-examine medical records from other deceased patients. Currently, many of the psychological symptoms associated with CTE are difficult to distinguish from other disorders, and brain-imaging tests for CTE are still in early development. Universal criteria for detecting CTE in donated brains could help scientists to search for other signatures so that physicians can diagnose the disease in living people. ■

CORRECTION

The News story 'Language origin debate rekindled' (*Nature* **518**, 284–285; 2015) misspelt the name of Paul Heggarty.

DWARF PLANETS: A TALE OF TWO MISSIONS

CALL IT THE YEAR of the dwarf planet. In 2015, scientists will get their first close-up look at two of the Solar System's biggest little rocks. The Dawn mission will fly past Ceres, in the asteroid belt between Mars and Jupiter, whereas New Horizons will encounter Pluto, the infamous ex-planet that orbits the icy reaches beyond Neptune. They promise to reveal surprises that could redefine how astronomers think of these small bodies.

MISSION COST

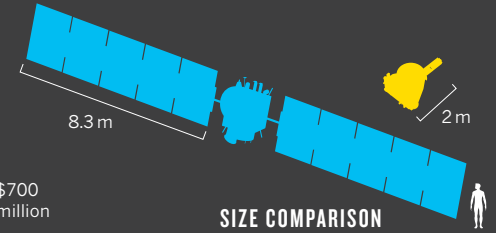
DAWN

US\$450 million

NEW HORIZONS

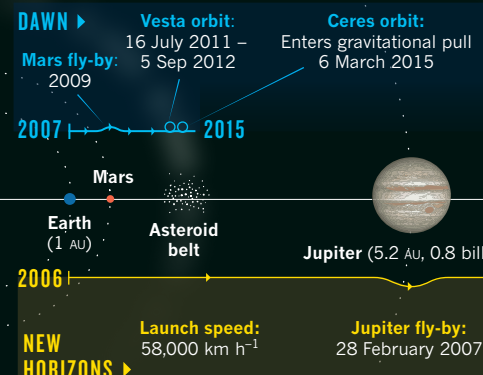
\$700 million

SIZE COMPARISON



GOING THE DISTANCE

The probes have been in space for similar lengths of time, but have covered vastly different distances.



DAWN TO CERES

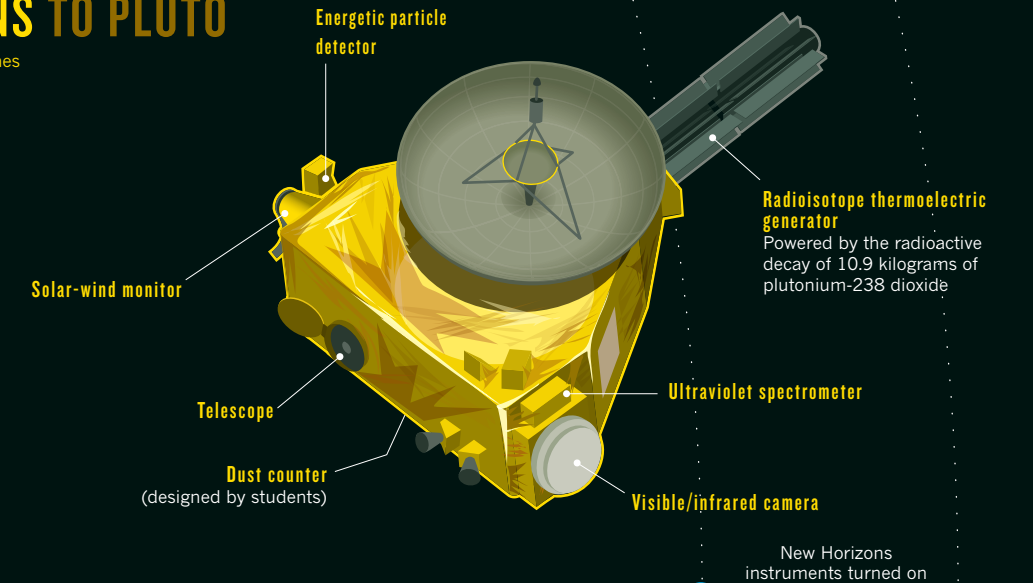
NASA's probe will analyse the largest unexplored objects in the inner Solar System.

LAUNCH: 27 SEPTEMBER 2007
TARGET: ASTEROID BELT

NEW HORIZONS TO PLUTO

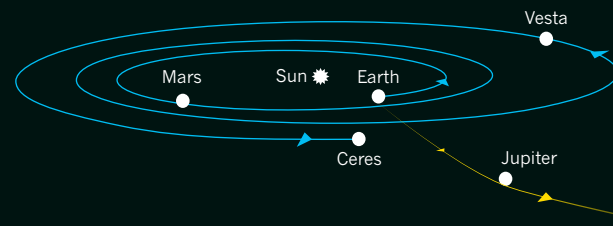
NASA's mission to the far reaches of the Solar System will gather data on a distant dwarf planet.

LAUNCH: 19 JANUARY 2006
TARGET: KUIPER BELT



DIFFERENT JOURNEYS

Dawn travelled using a mixture of thrusting and cruising on its way to the asteroid belt, whereas New Horizons blasted nearly directly outwards to Pluto.



New Horizons' speed received a boost of 14,400 km h⁻¹ from Jupiter to counter the gravitational pull of the Sun.

BODIES OF INTEREST

Both Ceres and Pluto are dwarf planets, but at first glance they have little in common.

Atmosphere: Thin, sporadic, watery

Temperature: -140°C to -70°C

Surface: Rocky, probably with buried ice

Key Ceres facts:

- Discovered in 1801 by Giuseppe Piazzi
- Largest object in the asteroid belt

Questions:

- How much of it is water?
- Was it once habitable?

WHAT'S IN A NAME?

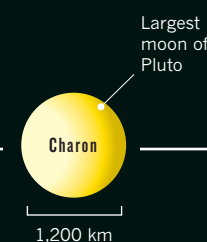
There is no faster way to trigger an argument among Solar System researchers than to bring up the definition of a planet. For decades, Pluto was considered the ninth planet. But in 2006, prompted by the discovery of other large Kuiper belt objects, the International Astronomical Union redefined what it means to be a planet. Pluto was declassified because it has not gravitationally cleared its orbit of other large bodies. Instead, Pluto and Ceres now belong to the newly created category of dwarf planets, which are allowed to orbit in a zone containing similar objects.

Key Pluto facts:

- Discovered in 1930 by Clyde Tombaugh
- First known Kuiper belt object

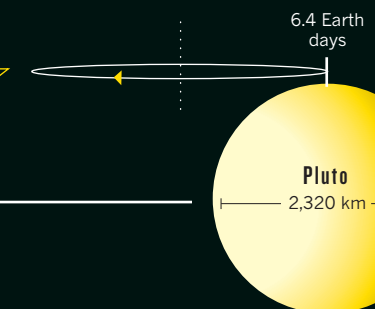
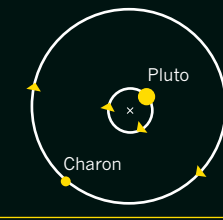
Questions:

- What does its icy surface look like?
- Was it ever geologically active?



Binary planets

Charon is so large compared to Pluto that the two both orbit a mutual centre of gravity, rather than one orbiting the other.



Atmosphere: Tenuous, probably replenished by ices sublimating from the surface

Temperature: -240°C to -220°C

Surface: Icy, with methane, nitrogen, carbon monoxide

BY ALEXANDRA WITZE / ILLUSTRATION BY NIK SPENCER

The Pluto siblings

Leslie and Eliot Young have spent their lives studying Pluto. Now they are gearing up for the biggest event of their careers.

BY ALEXANDRA WITZE



In a spare conference room in Boulder, Colorado, planetary scientists Leslie and Eliot Young quiz a graduate student to prepare him for his upcoming exams. They take their task

seriously, interrupting often as he answers questions about Pluto and Neptune's moon Triton.

Leslie makes a technical comment about the light reflecting off those distant worlds. Then, Eliot notes that Pluto and Triton may have started out very similar to one another in the early Solar System before evolving down different paths. "It's a classic case of nature versus nurture," he says. "They are siblings."

So, too, are the Youngs. Eliot and Leslie grew up as the oldest children of an astronautics researcher, and their mutual interests converged on one dwarf planet. "They are the only brother-sister Pluto team in the Solar System," says Alan Stern, a planetary scientist and principal investigator of NASA's New Horizons mission, which has been hurtling towards Pluto for the past nine years.

The Youngs and other Pluto researchers will be gearing up over the next few months as New Horizons finally nears its quarry, 4.8 billion kilometres from Earth. A telescope on the spacecraft has already begun capturing fuzzy pictures of Pluto, which will grow sharper as the probe closes in. And when New Horizons passes within 12,500 kilometres of Pluto on 14 July, it will provide the first close-up look at the world's icy surface, and the best chance yet to answer major questions about the evolution of the outer Solar System (see page 468).

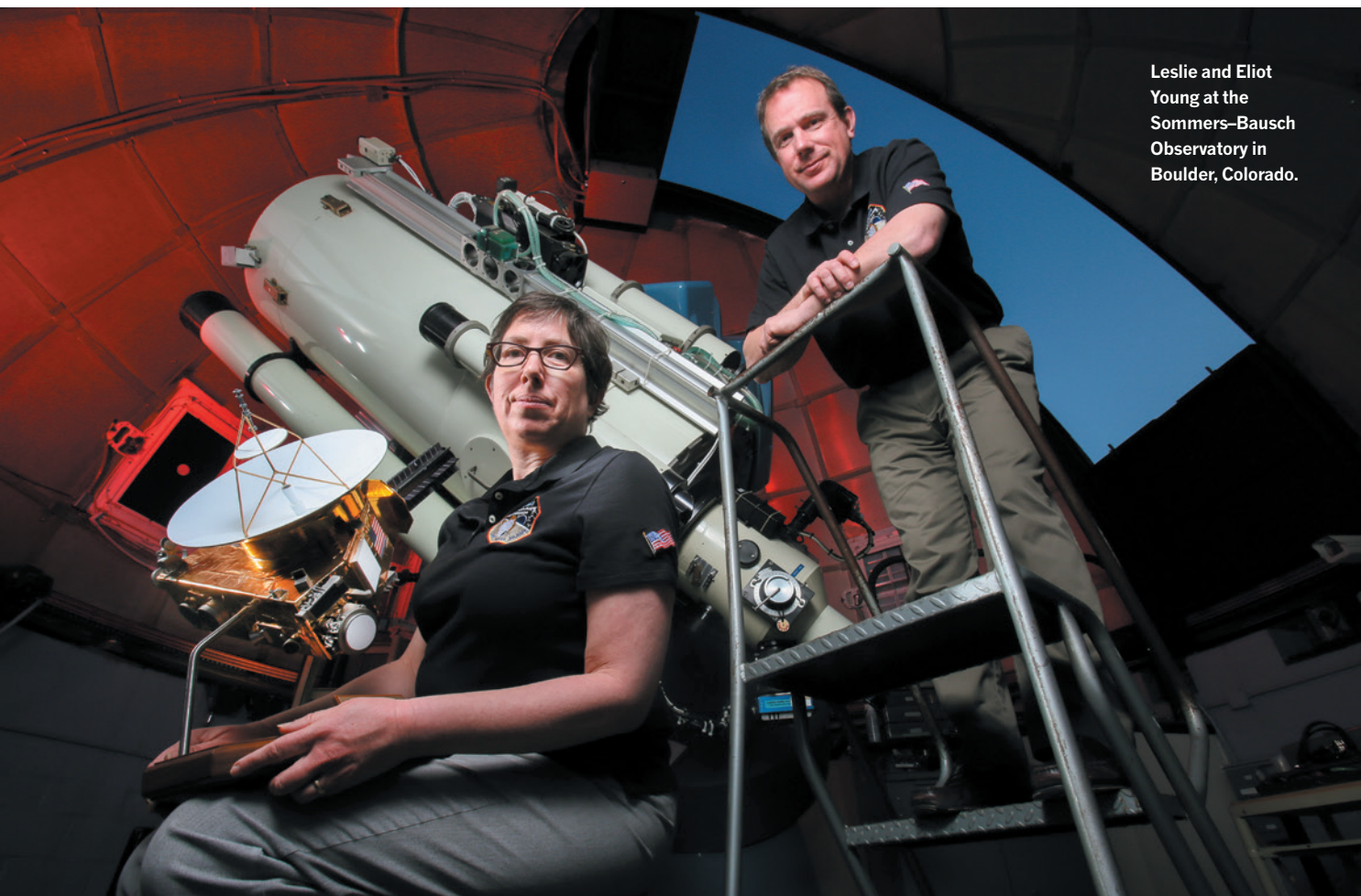
The fly-by will mark a major milestone in both the personal and the professional lives of the Youngs, who occupy adjoining offices at the Southwest Research Institute in Boulder. Over the past quarter of a century, their careers have intersected with Pluto science at key points, from helping to discover the dwarf planet's atmosphere to making some of the first detailed maps of its enigmatic surface. Whatever New Horizons finds this year will build in large part on work done by the siblings.

"We've had some ideas about how Pluto works for decades now," says Eliot. "We'll finally find out if they are right."

FAMILY ORBIT

When Eliot and Leslie were growing up in Newton, Massachusetts, family life revolved around their father, Larry Young, a legendary researcher at the Massachusetts Institute of Technology (MIT) in Cambridge. Young specializes in the biological effects of weightlessness, and he trained to fly on the space shuttle although he never went into orbit. Eliot, Leslie and their younger brother, Robert, sometimes played poker with visiting astronauts.

Larry Young was also a passionate skier who



Leslie and Eliot Young at the Sommers-Bausch Observatory in Boulder, Colorado.

BARRY GUTIERREZ

studied skiing injuries. On most winter weekends, his family took a long car trip to New Hampshire that was filled with brain games and chatter about mathematics.

Larry is not surprised that his two oldest children pursued science, but he never imagined them both studying the same dwarf planet in the distant reaches of the Solar System. “I think it’s the closest they could get to doing science fiction and still earn a living,” he says. (Robert ended up in software development.)

Eliot felt the pull of Pluto first. As a graduate student at MIT in the late 1980s, he worked with Jim Elliot and Ted Dunham, who were building instruments for airborne astronomy missions, including ones to study Pluto. But his sister, three years his junior, quickly followed. One day, she stopped in at his lab to show him a piece of computer coding she had done. Even though she was still an undergraduate at nearby Harvard University, Jim Elliot was impressed enough to offer her a job working on software.

The MIT team specialized in studying distant worlds using stellar occultations — when an object of interest moves between Earth and a background star. By measuring how much the light dims, planetary scientists can determine the size of the blocking

object. And by noting whether the light dims abruptly or gradually, they can deduce whether that object has an atmosphere. Pluto is so small (about two-thirds the size of Earth’s Moon) and far away (between about 30 and 50 times farther from the Sun than Earth is) that astronomers need to use every creative technique they can think of to tease out information.

SKYWATCH

One night in June 1988, several members of the MIT group took off from Honolulu, Hawaii, in the Kuiper Airborne Observatory, a telescope-carrying plane that flew above the obscuring effects of Earth’s atmosphere. Astronomers at the time suspected that Pluto had an atmosphere, but no one had ever spotted it.

Leslie Young was not yet a graduate student, but she was on the plane to help with the measurements. She distinctly remembers the excitement as the star’s light dimmed gradually, and Pluto’s long-sought atmosphere was revealed¹. The discovery, supported by ground-based measurements of the same occultation, made front-page headlines. When her brother Robert asked how it felt to have rewritten the textbooks, “I told him it felt pretty good,” she says.

Pluto remained a fuzzy dot on a map of the outer Solar System, even as other distant planets were coming into focus. NASA’s Voyager 2

spacecraft had visited Uranus in 1986, and three years later it swept past Neptune but did not go near Pluto. Researchers knew little about that world, other than that its surface seemed to be mostly ice, rather than rock, and it was accompanied by a moon, Charon, which was half the size of the dwarf planet itself. (Pluto was demoted from planet status in 2006.)

Helped by her keen coding skills, Leslie went on to make a series of major discoveries as part of the MIT team, including spotting methane in Pluto’s atmosphere² and nitrogen ice on its surface³. “If somebody asks me what’s my favourite colour,” she says, “I say 2.15 microns” — the wavelength of the light absorbed by the frozen nitrogen on Pluto’s surface.

In the years that followed, Leslie developed computer models to describe how the surface and atmosphere of Pluto interact. Because the orbit of the dwarf planet is extremely stretched out in an elongated ellipse, the amount of sunlight reaching its surface changes markedly throughout the Pluto year, which lasts 248 Earth years. When Pluto’s orbit carries it closer to the Sun, methane, nitrogen and other ices on the surface sublime and form a tenuous atmosphere, roughly one millionth the thickness of Earth’s. Some researchers argue that as Pluto gets farther away from the Sun in the coming years, the gases in the atmosphere

NATURE.COM
For more on Pluto,
see:
nature.com/pluto

will refreeze and drop to the surface, although Leslie's latest models suggest that the atmosphere never completely disappears⁴.

Occultation studies^{5,6} indicate that the density of Pluto's atmosphere doubled between 1988 and 2002 and has stayed pretty much constant since then. So one of New Horizons' major goals at Pluto is to unravel the icy interplay between the surface and the atmosphere.

FACE OF PLUTO

The big brother who got Leslie into Pluto has made his own mark. During and after his graduate studies, Eliot worked to map the face of Pluto by taking advantage of a geometric coincidence. Between 1985 and 1990, the orbital planes of Pluto and Charon tilted such that the two worlds regularly passed in front of one another as seen from Earth, allowing astronomers to watch a series of mutual eclipses. By measuring how Pluto's face dimmed as sections of it disappeared from view, researchers could work out which areas were dark and which were light — a property called albedo. Eliot was one of several scientists piecing together these mosaics to produce maps of the dwarf planet's surface⁷.

The maps were far from perfect: "The resolution is like somebody with a strong glasses prescription getting drunk and going to look at the Moon," says Eliot. But they provided some of the first real knowledge about what Pluto might look like. "It was foundational information that got people excited about Pluto," says Marc Buie, a Pluto astronomer now at the Southwest Research Institute. He competed with Eliot to generate the Pluto maps, and gives him credit for inventiveness. "He came up with some ways of tackling the data that I never would have thought of in a million years," says Buie.

Since then, the Hubble Space Telescope has managed to make sharper images of Pluto's surface. By around May this year, New Horizons will be close enough to Pluto to capture images better than Hubble's, and the dwarf planet will finally begin to come into focus. In the highest-resolution pictures, scientists should be able to pick out details as small as the lakes in New York City's Central Park.

Eliot is perhaps best known for his mapping work, but says that his most useful contribution to Pluto science is a method for modelling occultation light curves. He is happiest at the interface between technology and space science, especially if a healthy dose of hardware is involved. "I may still be an engineer at heart," he says. Dunham, now at Lowell Observatory in Flagstaff, Arizona, remembers Eliot building his own computers and housing them in cardboard boxes while in graduate school.

In recent years, Eliot has spent less time on Pluto and more on pushing technical boundaries in another frontier of Solar System science — sending balloons above Earth's atmosphere to make planetary observations. His propensity for technical tinkering has served him in his favourite hobby, too. He developed a new

timing system for races at a ski resort near Boulder, where he coaches a team.

Eliot and Leslie's mother, Jody Williams, is not surprised that the siblings have ended up working closely together. During their childhood trips to New Hampshire, where they had no television or other distractions, Eliot and



Eliot and Leslie with their parents in 1967.

Leslie would play for hours, building small towns and fantasy worlds with shared rules. Now, Williams says, "every time Eliot runs into a problem, he calls Leslie, and she never lets him down".

Today, Eliot retains a measure of his big-brother status: he is the gregarious one who often speaks for both of them and is known more widely among planetary astronomers. Leslie, more reserved, sometimes gets noticed initially because she is his sibling. "They learn I'm his sister and they figure I'm worth listening to," she says.

Yet the siblings clearly revel in their close relationship. Both usually work some mention of each other into professional discussions within minutes, and they constantly ricochet ideas between each other's offices. "They know each other better than any of us know either of them," says Stern.

FLY-BY FRENZY

As Pluto scientists approach peak excitement this year, Eliot will be helping to coordinate a cadre of amateur astronomers across the Southern Hemisphere who will try to capture a major occultation as Pluto passes in front of a particularly bright star on 29 June. It is the last of these events before the New Horizons fly-by, and a crucial data point in the series of occultation studies dating back to the mid-1980s — the only long-term measurements of

how Pluto's atmosphere has changed.

Leslie will move to mission control at the Johns Hopkins University Applied Physics Laboratory in Laurel, Maryland, to prepare for intense data gathering. She is a deputy project scientist for New Horizons and the chief architect of the details of the fly-by, a job that involves a steady diet of teleconferences and spreadsheets to coordinate what observation will be made by what instrument at what time. The mission team has carried out several dry runs of the entire encounter, trying to anticipate every possibility in what will be an action-packed few days leading up to and after the closest approach. "I've been working for the future for 15 years, and the pay-off is coming this summer," says Leslie.

Among other things, she has developed alternative trajectories for the spacecraft to divert to if it seems to be heading for a particularly dusty patch of space. With the spacecraft moving at nearly 50,000 kilometres per hour, collisions with dust particles could endanger it or damage instruments.

Cathy Olkin at the Southwest Research Institute, who was part of Jim Elliot's MIT group a few years after Leslie and is another New Horizons' deputy project scientist, says that observing occultations was good training for making high-stress, time-crucial measurements. She and Leslie have chased Pluto's shadow across islands in the Pacific and during snowstorms in New Zealand. "We know the value of telescope time and being prepared and having thought out what we're going to do at each step in time," says Olkin. "We know we have to get the data."

Back in her office, Leslie Young takes down her 1978 edition of David Halliday and Robert Resnick's *Fundamentals of Physics* textbook. She opens it to the back, to the reference table listing characteristics of Solar System objects. For Mercury, Venus, Earth and the rest of the planets, the table looks reassuringly full. For Pluto, the data column is incomplete.

One-third of the entries are question marks. Many of the rest are flat-out wrong. "Moons: none," it reads. (Charon had yet to be discovered.) "Atmosphere: none." (Ditto.)

Leslie runs her finger down the column, ticking off the Pluto discoveries that she and her big brother have been involved in. Atmosphere, radius, albedo, surface temperature — all key to understanding this curious little world. In July, she and Eliot hope to help fill in many of the remaining question marks. And maybe even add some new ones. ■

Alexandra Witze is a correspondent for *Nature* based in Boulder, Colorado.

1. Elliot, J. L. *et al. Icarus* **77**, 148–170 (1989).
2. Young, L. A. *et al. Icarus* **127**, 258–262 (1997).
3. Owen, T. C. *et al. Science* **261**, 745–748 (1993).
4. Olkin, C. B. *et al. Icarus* **246**, 220–225 (2015).
5. Elliot, J. L. *et al. Nature* **424**, 165–168 (2003).
6. Sicardy, B. *et al. Nature* **424**, 168–170 (2003).
7. Young, E. F. & Binzel, R. P. *Icarus* **102**, 134–149 (1993).



THE PAINFUL TRUTH

Brain-scanning techniques promise to give an objective measure of whether someone is in pain, but researchers question whether they are reliable enough for the courtroom.

BY SARA REARDON

Annie is lying down when she answers the phone; she is trying to recover from a rare trip out of the house. Moving around for an extended period leaves the 56-year-old exhausted and with excruciating pain shooting up her back to her shoulders. "It's really awful," she says. "You never get comfortable."

In 2011, Annie, whose name has been changed at the request of her lawyer, slipped and fell on a wet floor in a restaurant, injuring her back and head. The pain has never eased, and forced her to leave her job in retail.

Annie sued the restaurant, which has denied liability, for several

➤ NATURE.COM

For a podcast on measuring pain in the brain, see:

go.nature.com/ycdibz

hundred thousand dollars to cover medical bills and lost income. To bolster her case that she is in pain and not just malingering, Annie's lawyer suggested that she enlist the services of Millennium Magnetic Technologies (MMT), a Connecticut-based neuroimaging company that has a centre in Birmingham, Alabama, where Annie lives. MMT says that it can detect pain's signature using functional magnetic resonance imaging (fMRI), which measures and maps blood flow in the brain as a proxy for neural activity.

The scan is not cheap — about US\$4,500 — but Steven Levy, MMT's chief executive, says that it is a worthwhile investment: the company has had ten or so customers since it began offering the service in 2013, and all have settled out of court, he says. If the scans are admitted to Annie's trial, which is expected to take place early this year, it could establish a legal precedent in Alabama.

Most personal-injury cases settle out of court, so it is impossible to document how often brain scans for pain are being used in civil law. But the practice seems to be getting more common, at least in the United States, where health care is not covered by the government and personal-injury cases are frequent. Several companies have cropped up, and at least one university has offered the service.

The approach is based on burgeoning research that uses fMRI to understand the nature of pain — a very subjective experience. Scientists hope that the scans can provide an objective measure of that experience, and they see potential applications, such as in testing painkillers. But many neuroscientists say that the techniques are still far from being accurate enough for the courtroom. Critics say that the companies using them have not validated their tests or proved that they are impervious to deception or bias. And whereas some think the technologies will have a place in legal settings, others worry that the practice will lead to misuse of the scans.

"There's a real desire to come up with some more-objective proxy for pain," says Karen Davis, a neuroscientist at the University of Toronto in Canada. But such measures must be extremely accurate, she says. "The outcome of having a wrong answer can be quite catastrophic."

NEURAL ORIGINS

The methods that doctors commonly use to assess pain can seem crude. People are asked to rate their pain on a scale from one to ten, or choose from a row of cartoon faces that go from happy to anguished. These measures can help to chart changes in pain, as someone recovers from surgery, for example. But each person will experience and rate their pain differently, so one person's five could be worse than another's seven, and a nine might or might not be bad enough to keep someone from working.

An objective answer should lie in the brain, where the experience of pain is ultimately constructed. And although every experience is different, pain should share some common elements. Neuroscientist Tor Wager at the University of Colorado Boulder has been trying to decipher pain's signature in the brain by placing people in an fMRI scanner while they touch a hot plate. As the researchers turn the plate's temperature up and down, they record the activity across different parts of the brain, including the sensory regions associated with the hand. From these patterns, Wager says, they can predict with better than 90% accuracy whether the plate is just warm or painfully hot¹.

But this measures acute pain — the immediate response to an obvious stimulus. Chronic pain, like Annie's, affects hundreds of millions of people worldwide. And although its cause can be obvious, that is not always the case. Vania Apkarian of Northwestern University in Chicago, Illinois has scanned dozens of individuals soon after a back injury and then again over the course of a year or more. The pain went on to become chronic in roughly half of those people, and even though they described

the pain the same way throughout, Apkarian could detect a shift in the pain signature in their brains². It changed from a signal of activity in the insula, which is associated with acute pain, to one of activity in the medial prefrontal cortex, which processes cognitive behaviour, and the amygdala, which controls emotion. "Our interpretation is that the pain is becoming more internalized," Apkarian says.

This and other work suggests that there is an emotional component to chronic pain that is not necessarily involved in acute pain. Chronic pain and depression often coexist and reinforce one another. And some chronic pain can be eased with antidepressant drugs. But Wager cau-

tions that focusing on these links can be treacherous. Suggesting that pain is all in the head — even if that is technically the case — does not mean that it is imagined or faked. "People will always go to that black and white line," he says.

That line is a particular challenge in legal settings. "A person cannot be found disabled based on pain unless they can point to a specific cause," says Amanda Pustilnik, a legal expert at Harvard Law School in Cambridge, Massachusetts.

ISOLATED INSTANCES

The United States sees tens of thousands of injury lawsuits every year, most of which involve claims of unresolved pain.

But that might be unusually high — countries with national health systems, such as Canada, see fewer lawsuits, says Davis. So far, the only pain case involving brain-imaging techniques known to have progressed to trial involved a truck driver named Carl Koch, whose wrist was burned by a glob of molten asphalt in 2005. A year later, he said he was still in pain and sued his former employer, Western Emulsions in Tucson, Arizona, for damages.

Koch had had his brain scanned by Joy Hirsch, a neuroscientist who was running the fMRI Research Center at Columbia University in New York City. Hirsch had developed a method that she says can "tap into" chronic pain. Lightly touching the affected wrist provoked a signal in sensory regions and other brain areas associated with pain; touching the other wrist did not. The test, she says, is a well-characterized way to distinguish allodynia — a pain response to a stimulus that does not normally cause pain — from imagined pain.

At the trial, Western Emulsions called Sean Mackey, a neurologist at Stanford University in Redwood City, California, as an expert witness. Mackey maintained that pain is too subjective to measure in this way and that the signature Hirsch was detecting could have been produced if Koch had expected to feel pain in the affected wrist or was unduly concentrating on it — deliberately or not. Hirsch argued that there are known signals for imagined pain that were not apparent in the scans.

Ultimately, the judge admitted the scan, and the case settled for \$800,000 — more than ten times the company's initial offering, according to Koch's lawyer, Roger Strassburg.

Another issue, Mackey says, is that it might be possible for people to cheat the test. In a 2005 study, he instructed volunteers to lie in an fMRI scanner and touch a hot plate while he showed them a video of flames that became more or less intense on the basis of their brain activity. Given this visual feedback, volunteers were able to control the intensity of the flames by imagining the pain as being more or less severe than it actually was³. Mackey is looking into the technique as a way to control chronic pain, but he is also studying whether people can trick the scanner.

After the Koch case, the use of such techniques began to pick up. Hirsch, who is now at Yale University in New Haven, Connecticut, says that while she was at Columbia, she had been doing two to three pain-related scans per month, many of which were to support lawsuits. She is hoping to offer the service at Yale.

A main criticism of the various techniques being used in civil suits is

**"THERE'S A REAL DESIRE
TO COME UP WITH SOME
MORE-OBJECTIVE PROXY
FOR PAIN."**

the paucity of publications to validate them. Hirsch has not published anything on her method, but says that she does not think it is necessary. The way in which different body parts are represented in the brain has been well mapped, she says, and the scans she has done provide no further insight than answering whether or not the person was in pain.

MMT takes a somewhat different approach: it compares scans before and after an individual engages in a painful activity. For example, Annie was scanned before and after walking around, and the company claimed that it could detect a clear pain signal in the second scan. But the company's only publication, led by co-founder and chief science officer Donald Marks, has been a single case study. After the person did something painful, a brain scan revealed particularly strong activity in the insula, which is involved in consciousness and self-regulation, and the somatosensory cortex, which processes sensations from the various parts of the body⁴.

These regions are involved in pain, but they are also involved in many other things. "If you went to a Society for Neuroscience meeting and walked into any non-pain-related slide session, you'd see the same regions being talked about," Davis says. Getting a patient such as Annie to walk around between scans would not only cause her pain, but also increase her awareness of her back, which would activate the insula. Davis, who does not think that pain imaging should be used in court for this purpose, says that she finds it disturbing that Marks's study cites her work, which measured a different kind of brain activity. "It's quite shocking for them to be quoting studies that don't back up their technology at all," she says.

Moreover, the test cannot be validated in a single person, Wager says. Any number of confounding factors — emotion, expectation, or head movement in the scanner, for instance — could account for the signals the company sees. To prove that the method is valid, the researchers would have to show that the signals differ between people in pain and controls, he says, and that there is a biological mechanism that accounts for the signal. Without that, "it's like reading tea leaves."

Marks disputes this, saying that numerous studies, including Wager's, have shown that fMRI can reliably distinguish between pain states. "My work is an application on an individual basis of all the data to date which validates this approach," Marks says. He also argues that the approach is not meant to determine whether or not someone who says they are in pain actually is, "I'm taking individuals that everyone agrees have pain and providing a visual graphic representation of that pain."

CLOSE TO MARKET

Using different techniques, Chronic Pain Diagnostics (CPD) of Roseville, California, is planning to offer commercial scans for litigants. CPD compares scans taken of a person's brain after they received an electric shock to a database of images from 30 individuals with and without chronic pain. People with chronic pain respond to a stimulus differently from healthy controls, and the company has developed an algorithm that allows it to distinguish between the two with 92% accuracy⁵. CPD president and co-founder Shaun England says that he expects a scan to cost between \$5,000 and \$6,000.

Mackey says that the application is interesting and potentially useful if the technique is replicated in larger groups. But Apkarian says the sample size is too small to determine meaningful differences at this point. Just as in MMT's technique, background signals such as head movement could confound the interpretation. "If you simply blindly use it, there is a very good chance you will always find a difference" between groups, he says.

CPD's executive research director, Daniel Callan, says that the company

has ways to control for outside factors that could affect its database, such as randomizing the order in which the patients are scanned and using people of different ages and genders. But he agrees that further experiments are needed to determine how well the algorithm works for individual patients. England says that the company hopes to start another study soon.

Scientists' concerns about the validity of pain scans might not matter much to legal professionals and the courts, says Michael Flomenhaft, an attorney in New York City who specializes in chronic pain and neuroimaging. "There's a lot of scientific information that can't be stated with the level of certainty you'd need to present it at a scientific conference, but is confident and valuable in a legal setting."

There is, however, evidence that brain scans could be overly persuasive to jurors. Research has suggested that the general public is more likely to

accept poor arguments if they are accompanied by neuroscientific evidence⁶. In the Koch case, Mackey says, "pretty brain pictures ended up being very compelling".

The efforts to introduce pain imaging are similar in some ways to attempts over the past decade to use fMRI as a lie detector. Most researchers question the reliability of this technique. It is difficult to validate because study volunteers tend not to have the same motivations to lie as criminal defendants. But that has not stopped several companies from trying — thus far unsuccessfully — to have the evidence introduced in US courts. Pain imaging has been more successful owing to richer research on the topic. And the stakes are much lower for a civil case than in a criminal trial, so the bar for what constitutes evidence is lower, according to an analysis in the *Journal of Law and the Biosciences*⁷.

But some scientists and ethicists are concerned about where the increasing acceptance of pain imaging might lead. Pustilnik worries that it could become a sort of pass-fail test, not just forcing litigants to provide proof of their pain, but potentially making it a requirement to get prescription medications or insurance coverage. She is heading a working group at Harvard that is developing a list of ethical and scientific standards for the technologies before they become widespread.

Levy and Marks insist that their technology is not capable of that. "Fundamentally, we can't prove that a patient does not have pain," Levy says, because an individual might still be experiencing pain even if the scanner does not show it.

But that situation may be inevitable, says Stuart Derbyshire, a neuroscientist at the National University of Singapore. "If we accept the logic that the brain imager knows, then we have to accept that it's going to win even in cases when we don't want it to."

Even so, many say that the research should continue to strive for application, including inside the courtroom. "We already make many wrong treatment and legal decisions about who is and is not in pain and who shouldn't be believed," Wager says. "If we had new information, that could help us do a better job." ■ **SEE EDITORIAL P.456**

Sara Reardon writes for Nature from Washington DC.

1. Wager, T. D. et al. *N. Engl. J. Med.* **368**, 1388–1397 (2013).
2. Hashmi, J. A. et al. *Brain* **136**, 2751–2768 (2013).
3. deCharms, R. C. et al. *Proc. Natl Acad. Sci. USA* **102**, 18626–18631 (2005).
4. Marks, D. H., Valsasina, P., Rocca, M. A. & Filippi, M. *Internet J. Pain Symptom Contr. Palliat. Care* **8**, 1 (2009).
5. Callan, D., Mills, L., Nott, C., England, R. & England, S. *PLoS ONE* **9**, e98007 (2014).
6. Weisberg, D. S., Keil, F. C., Goodstein, S., Rawson, E. & Gray, J. R. *J. Cogn. Neurosci.* **20**, 470–477 (2008).
7. Salzmanowitz, N. *J. Law Biosci.* <http://dx.doi.org/10.1093/jlb/lsv003> (2015).

COMMENT

BIG DATA Three books on exabytes, in academia, business and governance **p.480**

DESERTS On the natural and cultural abundance of arid places **p.482**

REPRODUCIBILITY Validation is still vital, even if antibodies are recombinant **p.483**



OBITUARY Robert A. Berner, carbon-cycle geochemist, remembered **p.484**

SAMUEL ARANDA/PANOS



Nurses at the Kenema hospital in Sierra Leone, which contributed data to early efforts to sequence the genome of the Ebola virus from the West Africa outbreak.

Make outbreak research open access

Establish principles for rapid and responsible data sharing in epidemics, urge **Nathan L. Yozwiak, Stephen F. Schaffner and Pardis C. Sabeti.**

Last April, five months into the largest Ebola outbreak in history, an international group of researchers sequenced three viral genomes, sampled from patients in Guinea¹. The data were made public that same month. Two months later, our group at the Broad Institute in Cambridge, Massachusetts,

sequenced 99 more Ebola genomes, from patients at the Kenema Government Hospital in Sierra Leone.

We immediately uploaded the data to the public database GenBank (see go.nature.com/aotpbk). Our priority was to help curb the outbreak. Colleagues who had worked with us for a decade were at the

front lines and in immediate danger; some later died. We were amazed by the surge of collaboration that followed. Numerous experts from diverse disciplines, including drug and vaccine developers, contacted us. We also formed unexpected alliances — for instance, with a leading evolutionary virologist, who helped us to investigate ►

► when the strain of virus causing the current outbreak arose.

The genomic data confirmed that the virus had spread from Guinea to Sierra Leone, and indicated that the outbreak was being sustained by human-to-human transmission, not contact with bats or some other carrier. They also suggested new probable routes of infection and, importantly, revealed where and how fast mutations were occurring². This information is crucial to designing effective diagnostics, vaccines and antibody-based therapies.

What followed was three months of stasis, during which no new virus sequence information was made public (see 'Gaps in the data'). Some genomes are known to have been generated during this time from patients treated in the United States³. The number is likely to have been much larger: thousands of samples were transferred to researchers' freezers across the world.

In an increasingly connected world, rapid sequencing, combined with new ways to collect clinical and epidemiological data, could transform our response to outbreaks. But the power of these potentially massive data sets to combat epidemics will be realized only if the data are shared as widely and as quickly as possible. Currently, no good guidelines exist to ensure that this happens.

SPEED IS EVERYTHING

Researchers working on outbreaks — from Ebola to West Nile virus — must agree on standards and practices that promote and reward cooperation. If these protocols are endorsed internationally, the global research community will be able to share crucial

information immediately wherever and whenever an outbreak occurs.

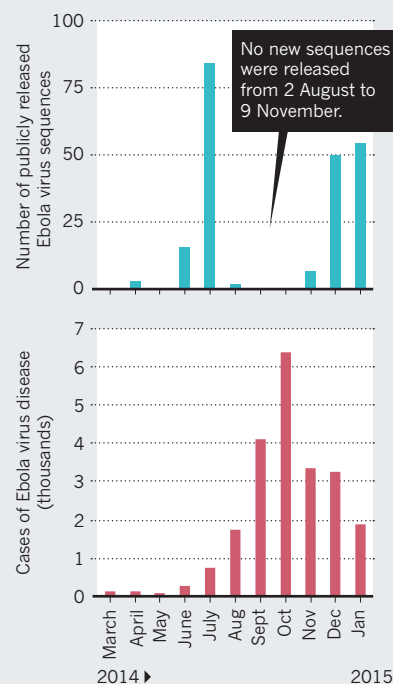
The rapid dissemination of results during outbreaks is sporadic at best. In the case of influenza, an international consortium of researchers called GISAID established a framework for good practice in 2006. Largely thanks to this, during the 2009 H1N1 influenza outbreak, the US National Center for Biotechnology Information created a public repository that became a go-to place for the community to deposit and locate H1N1 sequence information⁴. By contrast, the publishing of sequence information in the early stages of the 2012 Middle East respiratory syndrome (MERS) outbreak in Saudi Arabia highlighted uncertainties about intellectual-property rights, and the resulting disputes hampered subsequent access to samples.

Sharing data is especially important and especially difficult during an outbreak. Researchers are racing against the clock. Every outbreak can mobilize a different mixture of people — depending on the microbe and location involved — bringing together communities with different norms, in wildly different places. Uncertainties over whether the information belongs to local governments or data collectors present further barriers to sharing. So, too, does the absence of patient consent, common for data collected in emergencies — especially given the vulnerability of patients and their families to stigmatization and exploitation during outbreaks. Ebola survivors, for instance, risk being shunned because of fears that they will infect others.

Fortunately, useful models for responsible data sharing have been developed

GAPS IN THE DATA

Genome sequences from the West Africa outbreak of Ebola virus were first made publicly available in April 2014. Since 99 genomes were released in July, data sets have been shared sporadically, even though more are known to have been generated.



SOURCES: SEQUENCES, NCBI/VIROLOGICAL ORG; EBOLA CASES, WHO



Pilgrims in Saudi Arabia try to protect themselves from Middle East respiratory syndrome (MERS) virus.

by the broader genomics community. In 1996, at a summit held in Bermuda, the heads of the major labs involved in the Human Genome Project agreed to submit DNA sequence assemblies of 1,000 bases or more to GenBank within 24 hours of producing them^{5,6}. In exchange, the sequencing centres retained the right to be the first to publish findings based on their own complete data sets, by laying out their plans for analyses in 'marker' papers.

This rapid release of genomic data served the field well. New information on 30 disease genes, for instance, was published before the release of the complete human genome sequence. Since 1996, the Bermuda principles have been extended to other types of sequence data and to other fields that generate large data sets, such as metabolite research.

GUIDELINES FOR SHARING

More-recent policies on data release similarly seek to align the interests of different parties, including funding agencies, data producers, data users and analysts, and scientific publishers. Since January, for example, the US National Institutes of Health has required grantees to make large-scale genomics data public by the time of publication at the latest, with earlier deadlines for some kinds of data⁷.

HASAN JAMALI/AP

We urge those at the forefront of outbreak research to forge similar agreements, taking into account the unique circumstances of an outbreak.

First, incentives and safeguards should be created to encourage people to release their data quickly into the public domain. One possibility is to request that data users (and publishers) honour the publication intentions of data producers — the questions and analyses that they want to pursue themselves — for, say, six months. These intentions could be broadcast through several channels, including citable marker papers, disclaimer notices on data repositories such as GenBank, and online forums, such as virological.org and the EpiFlu database. Alternatively, data producers could publish an announcement about their data and their intentions on online forums as a resource that can be used by others as long as they cite the original source.

Second, ethical, rigorous and standardized protocols for the collection of samples and data from patients should be established to facilitate the generation and sharing of that information. A global consortium involving the leading health and research agencies and the ministries of health of engaged nations should work together towards establishing these. Ethicists should be involved to safeguard subjects' privacy and dignity. Biosecurity experts will also be needed to address potential dual-use research and other safety concerns. A helpful analogue is the approach used by the Human Heredity and Health in Africa (H3Africa) Initiative, which aims to apply genomics to improving the health of African populations. Since August 2013, H3Africa has used standard consent-form guidelines⁸ for collecting DNA samples from subjects for genomic studies, regardless of their country of origin.

Lastly, any preparation for future outbreaks should include provisions for rapidly building new bridges and establishing community norms. Successful collaborations in genomics and historical data-sharing agreements have tended to involve a fairly stable group of individuals and organizations, making norms of behaviour relatively easy to establish and sustain. By contrast, outbreaks can involve a new cast of characters each time, and cases in which the pathogen is new to science call for whole new fields of research.

THE KENEMA WAY

As a first step, we call on health agencies such as the World Health Organization, the US Centers for Disease Control and Prevention and Médecins Sans Frontières, as well as genome-sequencing centres and other research institutions, to convene a meeting this year — similar to that held in Bermuda



Quarantine officers rush to test passengers at Tokyo's Narita airport amid the 2009 swine-flu outbreak.

in 1996. Attendees must include scientists, funders, ethicists, biosecurity experts, social scientists and journal editors.

We urge researchers working on outbreaks to embrace a culture of openness. For our part, we have released all our sequence data as soon as it has been generated, including that from several hundred more Ebola samples we recently received from Kenema. We have listed the research questions that we are pursuing at virological.org and through GenBank, and we plan to present our results at virological.org as we generate them, for others to weigh in on. We invite people either to join our publication, or to prepare their own while openly laying out their intentions online. We have also made clinical data for 100 patients publicly available and have incorporated these into a user-friendly data-visualization tool, Mirador, to allow others to explore the data and uncover new insights.

"We urge researchers working on outbreaks to embrace a culture of openness."

Kenema means 'translucent, clear like a river stream' or 'open to the public gaze'. To honour the memory of our colleagues who died at the forefront of the Ebola outbreak, and to ensure that no future epidemic is as devastating, let's work openly in outbreaks. ■

Nathan L. Yozwiak and Stephen F. Schaffner are senior staff scientists, and **Pardis C. Sabeti** is associate professor, at the Broad Institute and Harvard University in Cambridge, Massachusetts, USA. e-mail: nyoziwak@broadinstitute.org

1. Baize, S. *et al.* *N. Engl. J. Med.* **371**, 1418–1425 (2014).
2. Gire, S. K. *et al.* *Science* **345**, 1369–1372 (2014).
3. Vogel, G. *Science* **346**, 684–685 (2014).
4. Bogner, P. *et al.* *Nature* **442**, 981 (2006).
5. Marshall, E. *Science* **291**, 1192 (2001).
6. Toronto International Data Release Workshop. *Nature* **461**, 168–170 (2009).
7. US National Institutes of Health. 'NIH Genomic Data Sharing Policy' (27 August 2014); available at <http://go.nature.com/i88ot5>
8. H3 Africa. *H3 Africa Guidelines for Informed Consent* (H3Africa, 2013).
9. Jackson, M. D. *Life Within Limits: Well-being in a World of Want* (Duke Univ. Press, 2011).

BIG DATA

The revolution is digitized

Charles Seife digs into three studies of the wild new world of big data.

The term has been around for almost two decades, but the world only really started talking about 'big data' in the first few months of 2011. We know this because we can look it up on Google Trends (see 'Byte marks').

Google built its empire on gathering and analysing nearly unfathomable depths of data. Every query ever typed into the search engine is sitting in Google's multi-exabyte data stores. These stores also hold the full text of tens of millions of books, high-resolution images of streets around the world and myriad e-mails, videos, word-processing documents and spreadsheets. Anything that can be rendered in bits and bytes and is accessible to the company's servers will be pushed, filed, stamped, indexed, briefed, debriefed and numbered by semi-autonomous information-gathering agents. Enter 'big data' into the Google Trends website and, a fraction of a second later, a graph of frequency appears, its line rising sharply upwards in the first quarter of 2011. You are distilling that information from a colossal data set containing the entire world's search-engine queries for the past ten years.

Seamless upgrades in computer interfaces masked a liminal moment: in a few years, we have moved from data that can be created, gathered and understood by unaided humans — kilobytes, megabytes and gigabytes — into the hitherto unimaginable realm of petabytes and exabytes, gathered at terahertz speeds and processed almost as quickly. The transition has moved beyond scale to revolution.

In *Big Data, Little Data, No Data*, information-studies specialist Christine Borgman looks at big data through a fairly narrow lens: academic research. Each day, scientists grapple with ever more appalling volumes of data. The ATLAS detector on the Large Hadron Collider at CERN, Europe's particle-physics laboratory near Geneva, Switzerland, has to sort through dozens of terabytes of data every second while it is running — and filter that down by five orders of magnitude before humans can deal with it. Next-generation

Big Data, Little Data, No Data: Scholarship in the Networked World

CHRISTINE L. BORGMAN
MIT Press: 2015.

Data-ism: Inside the Big Data Revolution

STEVE LOHR
Oneworld: 2015.

Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World

BRUCE SCHNEIER
W. W. Norton: 2015.

telescopes such as the Square Kilometre Array will be gathering exabytes of data each day — an amount that would have filled the total storage capacity of all the world's information-carrying devices (including books, photos and videos) up to the mid-1980s.

Borgman is something of a data anthropologist. She goes among researchers in physical sciences, social sciences and humanities alike to find out how they collect, handle and share the flood of information. Her treatise is interesting, but frustrating. She has difficulty turning her sizeable data set into a narrative both broad enough to cover the range of topics and deep enough to do justice to them. All too often, she seems to give a quick nod to essential elements. For example, she mentions open publications and data, but provides no hint of the battles around them in the research and publishing worlds. She offers key insights — that there are different dynamics to publishing research results and raw data, and that it is shortsighted to focus on releasing new data sets rather than on how to preserve and reuse the data. But the book might have said much more. There is nary a word about the huge controversy around incursions of commercial entities into the gathering, dissemination and control of scholarly data.

In *Data-ism*, Steve Lohr goes after the commercial implications of big data, but through an equally narrow lens. As a veteran technology and business reporter, he is attracted to the story of how data can help to root out

BYTE MARKS

The surge of interest in big data since 2011 can be clearly traced in Google's archive of search terms.

SOURCE: GOOGLE TRENDS; GRAPHIC: CLAIRE WELSH/NATURE

inefficiencies that stop businesses reaching their potential. He gives the example of McKesson, a drug and medical-supply distributor that used its archives of product and shipping data to create a supply-chain model. That led to a billion-dollar decrease in inventories and a sizeable jump in efficiency, showing, as Lohr says, “data really being used to ... make better decisions, ones that trump best guesses and gut feel, experience and intuition”.

Alas, *Data-ism* is very much a conventional business book, full of anecdotes, mini-profiles and aphorisms that grow ever less compelling, however well they would go over at a TEDx talk. Lohr’s journalistic instincts often seem to betray him. He is unimpressed with the massive data-collecting and consumer-profiling of information giant Acxiom, yet bowled over by a seemingly conventional personality-horoscope program that snaffled up Twitter feeds, and, for 81% of subjects, “pretty much matched the results of their formal tests for personality type, basic values, and needs”.

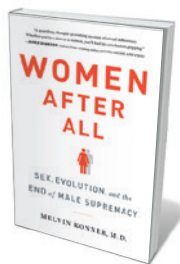
Neither Borgman nor Lohr truly grapples with the immensity of the big-data story. At its core, big data is not primarily a business or research revolution, but a social one. In the past decade, we have allowed machines to act as intermediaries in almost every aspect of our existence. When we communicate with friends, entertain ourselves, drive, exercise, go to the doctor, read a book — a computer transmitting data is there. We leave behind a vast cloud of bits and bytes.

Bruce Schneier, a security analyst known for designing the Blowfish block-cipher algorithm — a fast and flexible method of encrypting data — grasps this revolution’s true dimensions. In *Data and Goliath*, he describes how our relationships with government, corporations and each other are transformed by ordinary, once-ephemeral human interactions being stored in digital media. The seemingly meaningless, incidental bits of data that we shed are turning the concept of privacy into an archaism, despite half-hearted (and doomed) regulations to protect “personally identifiable information”. As science-fiction pioneer Isaac Asimov wrote some 30 years ago: “Things just seem secret because people don’t remember. If you can recall every remark, every comment, every stray word made to you or in your hearing and consider them all in combination, you find that everyone gives himself away in everything.”

Schneier paints a picture of the big-data revolution that is dark, but compelling; one in which the conveniences of our digitized world have devalued privacy. Interest in privacy has dropped by 50% over the past decade — at least according to Google Trends. ■

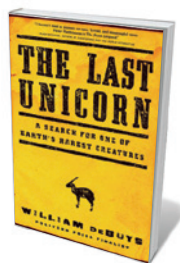
Charles Seife is a professor of journalism at New York University, and the author of books including *Proofiness* and *Virtual Unreality*. e-mail: cs129@nyu.edu

Books in brief



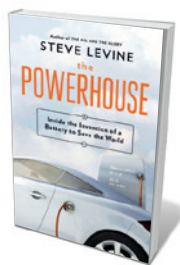
Women After All: Sex, Evolution, and the End of Male Supremacy Melvin Konner W. W. NORTON (2015)

The mammalian body plan is basically female, and maleness is a syndrome. So declares anthropologist Melvin Konner in this biologically based study (although recent research points to complexities; see *Nature* **518**, 288–291; 2015). Positing that women are more altruistic and pragmatic — and so are best-equipped for the future — Konner mines evolution and anthropology to probe gender identities in the light of biology, sexual conflict across species and more. The provocative scenarios he lays out include a man-free world where women reproduce using DNA from other women’s eggs.



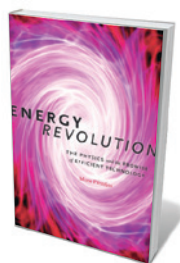
The Last Unicorn: A Search for One of Earth's Rarest Creatures William deBuys LITTLE, BROWN (2015)

Discovered in 1992, the saola (*Pseudoryx nghetinhensis*) is one of the rarest large mammals, a beautiful ruminant found in the mountains between Laos and Vietnam. In 2011, nature writer William deBuys and field biologist William Robichaud set out to gauge poaching pressures on the saola. DeBuys’ account of destitute villages and endangered animals left to die in snares is a familiar narrative of conservation in poor countries. But, like Peter Matthiessen’s 1978 *The Snow Leopard* (Viking), this is less an homage to an iconic species than a meditation on our compulsion to harry and hem in the wild.



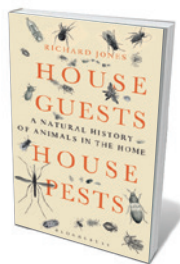
The Powerhouse: Inside the Invention of a Battery to Save the World Steve Levine VIKING (2015)

Journalist Steve Levine’s chronicle of the race to develop a rechargeable lithium-ion electric-car battery makes for a propulsive techno-saga. The action centres on the Argonne National Laboratory outside Chicago, Illinois, where an international group led by engineer Jeff Chamberlain worked on the knotty physics. Levine interweaves the geopolitical jostling of the US lab and others in Asia, climaxing with Argonne’s 2012 win of more than US\$120 million to build the ‘Hub’ — a powerhouse intended to create a sustainable battery industry.



Energy Revolution: The Physics and the Promise of Efficient Technology Mara Prentiss HARVARD UNIVERSITY PRESS (2015)

In this crisp, evidence-based treatise, physicist Mara Prentiss makes a remarkable assertion: that solar and wind power could supply 100% of average US energy needs for the next 50 years. Prentiss argues that a transition to renewables is probable, given that energy revolutions are a historical norm. She stacks up reams of salient data, such as the fact that US energy use per capita has remained steady since 1965, thanks to increasing fuel efficiency. Although optimistic, her analyses of energy sources, combinations, conservation and storage compel.



House Guests, House Pests: A Natural History of Animals in the Home Richard Jones BLOOMSBURY (2015)

Urban nature lovers relish the sight of birds or hedgehogs in their gardens. “Something odd, though, happens at the back door,” notes Richard Jones — and that is zero tolerance for wild unbidden guests, from tapestry moths to rats. Jones, a fellow of the Royal Entomological Society, is a learned guide to this alarming panoply of intruders, from the bacon beetle (*Dermestes lardarius*), a vagrant of old-fashioned larders, to the noisy edible dormouse (*Glis glis*), which can infest the attics of rural houses. *Barbara Kiser*

Correspondence

Gear students up for big medical data

Combining big data with personalized medicine is an unprecedented opportunity. It will probably be cheaper than current practices in the long term, particularly given the questionable effectiveness of many medications (see *Nature* **517**, 540; 2015).

Success in this endeavour will depend on training the next generation of clinicians and data scientists to deploy terabytes of data to select from a range of diagnosis and treatment options.

Undergraduate and graduate bioinformatics programmes need to embrace data-analytics courses geared towards generating a new type of medical specialist — one who no longer needs to see patients, just their data.

Ervin Sejdić *University of Pittsburgh, Pennsylvania, USA.*
esejdic@ieee.org

Antibodies: the solution is validation

I disagree with Andrew Bradbury and colleagues' suggestion that making the sequences of commercial antibodies publicly available could minimize irreproducibility in biomedical research (*Nature* **518**, 27–29; 2015). The real solution is proper initial validation of antibodies.

In my view, the reproducibility problem is better addressed by identifying the good antibodies and the reputable companies that develop, validate and manufacture them — as astute scientists do now. Also, journals need to mandate the provision of detailed validation data, protocols and antibody sources (clone, catalogue number). Independent websites enabling the submission of antibody data and consumer feedback would also help.

The biggest investment in developing a good monoclonal antibody is the extensive work

needed to validate specificity and sensitivity across all relevant applications. Unlike therapeutic antibodies, most research antibodies are not sequence-patented because the cost is too high to be recovered by sales.

Even if the practical hurdles of funding and enforcing a sequence-publishing policy could be overcome, making unpatented antibody sequences public would allow them to be widely copied, produced and sold. This would eliminate the incentive for good companies to invest in validation. It would also allow 'bad' antibody sequences to contaminate the databases.

The authors' proposal could therefore disproportionately harm the good companies, hurt the end-users it is designed to protect, and would not solve the reproducibility problem.

Roberto D. Polakiewicz *Cell Signaling Technology, Danvers, Massachusetts, USA.*
rpolakiewicz@cellsignal.com

Antibodies: validate recombinants too

Recombinant antibodies are pure proteins with minimal batch-to-batch variability, so could provide an important element of antibody standardization (A. Bradbury *et al.* *Nature* **518**, 27–29; 2015). However, they must still be functionally validated if they are to help solve the reproducibility crisis.

Vendors and researchers would have to optimize their recombinant antibodies for specific applications, because of the inherent complexity of these molecules and their ability to bind non-specifically to other proteins carrying similar immunological sequences.

Suppliers and users of such antibodies will need specialized training in this validation and optimization, particularly in experimental design and the extensive use of controls.

Leonard P. Freedman *Global Biological Standards Institute,*

Washington DC, USA.
lfreedman@gbsi.org

Polluters migrate to China's poor areas

Bo Zhang and Cong Cao argue that China's citizens should have a legal right to safeguard the quality of their environment (*Nature* **517**, 433–434; 2015). The wealthy would stand to benefit most from such a public litigation system, causing pollution producers to migrate to poorer areas.

Heavy industry in China is already moving out of developed eastern regions to the west (see X. Bai *et al.* *Nature* **509**, 158–160; 2014), where it is damaging the local ecology. Industrial waste slag has eroded a nature conservation area in Xinjiang (see go.nature.com/68e1lo; in Chinese), for example, and discharge from factories has severely polluted part of the Tengger Desert at the border of the Inner Mongolia and Ningxia regions (see go.nature.com/nlfbrs; in Chinese).

Environmental activism by residents in affluent areas such as Shenzhen, Jinan and Beijing (see, for example, Q. Wang *Nature* **497**, 159; 2013) is accelerating this migration of polluters into poor areas where environmental protection is considered a luxury, and where water and soil are already badly contaminated.

Xin Miao *Harbin Institute of Technology, Harbin, China.*

Yanhong Tang *Northeast Agricultural University, Harbin, China.*

Christina W. Y. Wong *The Hong Kong Polytechnic University, Kowloon, Hong Kong.*
xin.miao@aliyun.com

Biochar: bring on the sewage

Biochars are carbon-rich soil additives derived from agricultural and other plant waste that could enhance crop productivity (see *Nature* **517**, 258–260; 2015). We suggest that

biochars could also be produced from human sewage — an underutilized resource that is rich in soil nutrients and carbon.

Sanitation problems in developing regions would be alleviated by diverting sewage solids into producing biochar, made by thermal conversion in sealed containers. This might even offset the need to install conventional sewage-treatment infrastructure with its higher construction and operation costs.

Marc Breulmann, Manfred van Afferden, Christoph Fühner *Helmholtz Centre for Environmental Research — UFZ, Leipzig, Germany.*
christoph.fuehner@ufz.de

Biochar: pros must outweigh cons

To optimize the agricultural and environmental benefits of biochar, a charcoal-rich soil additive, we need to overcome its potentially undesirable effects (see *Nature* **517**, 258–260; 2015).

For example, it is uncertain whether biochar — effectively an underground carbon store — can help to mitigate carbon emissions. A ten-year study of boreal forests found that applying biochar led to soil degradation and increased the activity of soil microbes, causing carbon dioxide release (D. A. Wardle *et al.* *Science* **320**, 629; 2008).

Adding blackened biochar can also lower the reflectivity (albedo) of the soil surface, potentially exacerbating climate warming (S. Meyer *et al.* *Environ. Sci. Technol.* **46**, 12726–12734; 2012).

Tilling deep furrows in the soil would help to reduce the decline in reflectivity and increase the efficiency of applied biochar.

However, this practice could also encourage carbon dioxide release.

Hong Yang *University of Oslo, Norway.*

Xianjin Huang *Nanjing University, China.*

Julian R. Thompson *University College London, UK.*
hongyanghy@gmail.com

Robert A. Berner

(1935–2015)

Geochemist who quantified the carbon cycle.

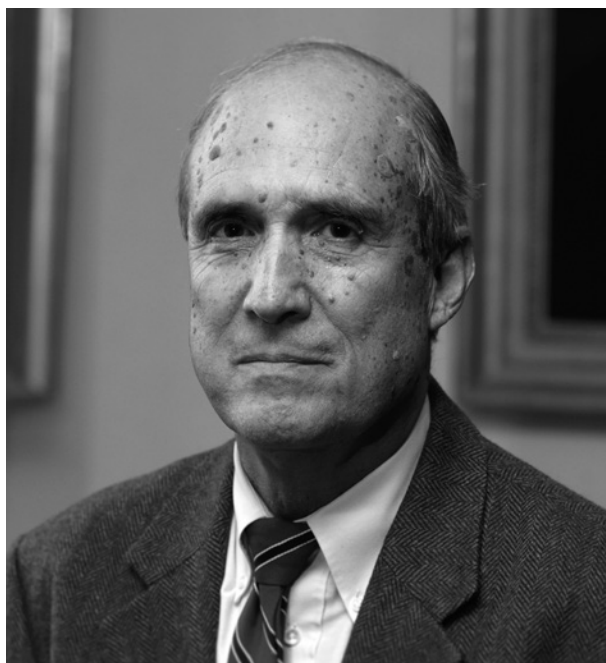
From how minerals form in sediments to how carbon dioxide is regulated in the atmosphere, Robert Arbuckle Berner quantified elemental cycles across the Earth system. He developed the first whole-Earth mathematical model of CO₂ exchange, which revealed marked changes in our planet's past atmospheric levels and the rates at which natural processes might remove anthropogenic CO₂ from the atmosphere.

Born in 1935 in Erie, Pennsylvania, Berner died on 10 January in New Haven, Connecticut. He was encouraged to develop an interest in geology by his older brother Paul, a (now-retired) petroleum geologist. Berner attended the University of Michigan in Ann Arbor for his undergraduate and master's degrees. There, he spotted fellow geology student Betty Kay. They married in 1959 and formed an inseparable bond, working and writing papers and books together for decades.

Berner received his PhD in 1962 from Harvard University in Cambridge, Massachusetts. During his thesis work on the formation of iron sulfides in sediments, he discovered new minerals, among them greigite, and invented a type of electrode used for measuring sulfide content. His adviser was Raymond Siever, known for his work on the ancient marine silicon cycle, and Berner was also heavily influenced by Bob Garrels, who championed thermodynamics and the concept of geochemical cycles.

He moved as a postdoc to the Scripps Institution of Oceanography in La Jolla, California. After a short stay as assistant professor at the University of Chicago, Illinois, in 1965 he joined the faculty at Yale University in New Haven, where he remained until his retirement in 2006.

Soon after arriving at Yale, Berner realized that mineral formation in sediments depends on how fast chemicals are transported in and out of the sediments and how quickly organic matter is oxidized by microbes. He developed mathematical expressions for these mechanisms and so started the field of sediment diagenesis, which concerns the biological and chemical processes that occur in recently formed



sediments. Berner and others went on to establish how sediment processes ultimately control the nutrient balance of the oceans and the concentrations of oxygen and CO₂ in the atmosphere.

In the early 1980s, Berner teamed up with Garrels and Antonio Lasaga to develop the BLAG (Berner, Lasaga and Garrels) model of global atmospheric CO₂ concentrations over geological time. This was the first global model aimed at quantifying all conceivable processes that control CO₂ exchange, and was largely based on Berner's earlier work on mineral-weathering reactions, ocean chemistry and early diagenesis. The BLAG model allowed geologists to understand for the first time how changes in rates of geological processes such as continental plate motion, for example, controlled past CO₂ levels.

Components since added to the model include the influence of biological evolution on the history of CO₂ concentrations, which hint at relationships between plant evolution and glaciation. These later models also reproduce a history of atmospheric oxygen, and show, for instance, how past periods of elevated oxygen concentrations correlate with spells of insect gigantism.

Bob focused like a laser beam on the problem at hand and was able to find simple and elegant solutions to complex

geological problems. He exuded warmth and humanity. In my time at Yale as his PhD student, from 1982 to 1988, he often joined us for Friday happy hour at the local Whitney Winery, entertaining us with stories of science's colourful characters.

He was a Francophile and loved the winery's outside terrace because it reminded him of Parisian cafés. It was more expensive to sit there so Bob inevitably picked up the bill, and sometimes invited us home afterwards for a meal and to sample his (not so fine) wines and whiskey. Fine wine was saved for the celebration of new PhDs. Bob and Betty invited the newly minted PhD to their home for a luxury dinner. Afterwards, the entire lab would descend, often for a long night of ping-pong and poker.

In material things, Bob had simple tastes. When he and Betty inherited a powder-blue Chevy Nova, Bob gleefully announced that it was his first car with a radio, and invited a group of graduate students to ride with him through New Haven listening to old-time radio stations. Bob later purchased a Honda Civic, which he named Harvey, that had a tape deck and door alarm. He made a 4-track tape recording of piano and percussion parts, which, when played in the car with front doors ajar, mixed perfectly with the car alarm. Bob had a party to celebrate, packing graduate students into the car to enjoy his composition, 'Harvey and the four Bobs'.

He was, however, a serious pianist and classical composer, devoting much time to music after his retirement (see go.nature.com/bnq4a2).

Bob was a loving family man and a dedicated friend and mentor. He stressed honesty and integrity while showing that science was great fun. Bob touched the whole geological community. He is sorely missed. ■

Don Canfield is professor of ecology and director of the Nordic Center for Earth Evolution (NordCEE) at the University of Southern Denmark in Odense, Denmark. He was a PhD student of Bob Berner's from 1982 to 1988 at Yale University in New Haven, Connecticut.
e-mail: dec@biology.sdu.dk

YALE UNIV.

ARTIFICIAL INTELLIGENCE

Learning to see and act

An artificial-intelligence system uses machine learning from massive training sets to teach itself to play 49 classic computer games, demonstrating that it can adapt to a variety of tasks. [SEE LETTER P.529](#)

BERNHARD SCHÖLKOPF

Improvements in our ability to process large amounts of data have led to progress in many areas of science, not least artificial intelligence (AI). With advances in machine learning has come the development of machines that can learn intelligent behaviour directly from data, rather than being explicitly programmed to exhibit such behaviour. For instance, the advent of 'big data' has resulted in systems that can recognize objects or sounds with considerable precision. On page 529 of this issue, Mnih *et al.*¹ describe an agent that uses large data sets to teach itself how to play 49 classic Atari 2600 computer games by looking at the pixels and learning actions that increase the game score. It beat a professional games player in many instances — a remarkable example of the progress being made in AI.

In machine learning, systems are trained to infer patterns from observational data. A particularly simple type of pattern, a mapping between input and output, can be learnt through a process called supervised learning. A supervised-learning system is given training data consisting of example inputs and the corresponding outputs, and comes up with a

model to explain those data (a process called function approximation). It does this by choosing from a class of model specified by the system's designer. Designing this class is an art: its size and complexity should reflect the amount of training data available, and its content should reflect 'prior knowledge' that the designer of the system considers useful for the problem at hand. If all this is done well, the inferred model will then apply not only for the training set, but also for other data that adhere to the same underlying pattern.

The rapid growth of data sets means that machine learning can now use complex model classes and tackle highly non-trivial inference problems. Such problems are usually characterized by several factors: the data are multidimensional; the underlying pattern is complex (for instance, it might be nonlinear or changeable); and the designer has only weak prior knowledge about the problem — in particular, a mechanistic understanding is lacking.

The human brain repeatedly solves non-trivial inference problems as we go about our daily lives, interpreting high-dimensional sensory data to determine how best to control all the muscles of the body. Simple supervised learning is clearly not the whole story, because

we often learn without a 'supervisor' telling us the outputs of a hypothetical input–output function. Here, 'reinforcement' has a central role in learning behaviours from weaker supervision. Machine learning adopted this idea to develop reinforcement-learning algorithms, in which supervision takes the form of a numerical reward signal², and the goal is for the system to learn a policy that, given the current state, determines which action to pick to maximize an accumulated future reward.

Mnih *et al.* use a form of reinforcement learning known as Q-learning³ to teach systems to play a set of 49 vintage video games, learning how to increase the game score as a numerical reward. In Q-learning, $Q^*(s,a)$ represents the accumulated future reward, Q^* , if in state s the system first performs action a , and subsequently follows an optimal policy. The system tries to approximate Q^* by using an artificial neural network — a function approximator loosely inspired by biological neural networks — called a deep Q-network (DQN). The DQN's input (the pixels from four consecutive game screens) is processed by connected 'hidden' layers of computations, which extract more and more specialized visual features to help approximate the complex

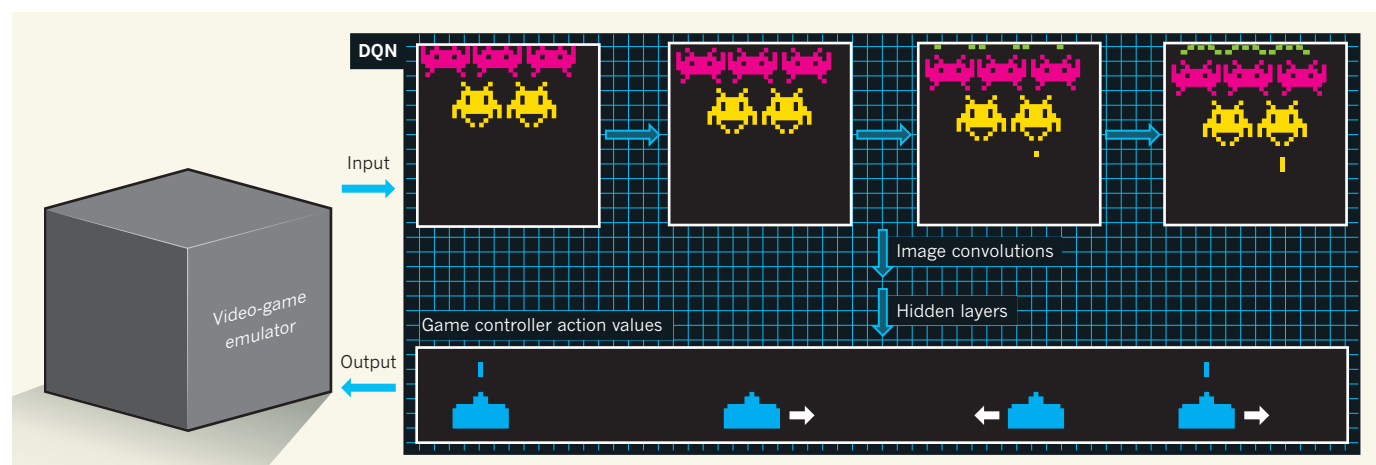


Figure 1 | Computer gamer. Mnih *et al.*¹ have designed an artificial-intelligence system, using a 'deep Q-network' (DQN), that learns how to play 49 video games. The DQN analyses a sequence of four game screens simultaneously and approximates, for each possible action it can make, the consequences on the future game score if that action is taken and followed by the best possible course of subsequent actions. The first layers of the DQN

analyse the pixels of the game screen and extract information from more and more specialized visual features (image convolutions). Subsequent, fully connected hidden layers predict the value of actions from these features. The last layer is the output — the action taken by the DQN. The possible outputs depend on the specific game the system is playing; everything else is the same in each of the 49 games.

nonlinear mapping between inputs and the value of possible actions — for instance, the value of a move in each possible direction when playing *Space Invaders* (Fig. 1).

The system picks output actions on the basis of its current estimate of Q^* , thereby exploiting its knowledge of a game's reward structure, and intersperses the predicted best action with random actions to explore uncharted territory. The game then responds with the next game screen and a reward signal equal to the change in the game score. Periodically, the network uses inputs and rewards to update the DQN parameters, attempting to move closer to Q^* . Much thought went into how exactly to do this, given that the agent collects its own training data over time. As such, the data are not independent from a statistical point of view, implying that most of statistical theory does not apply. The authors store past experiences in the system's memory and subsequently re-train on them — a procedure they liken to hippocampal processes during sleep. They also report that the system benefits from randomly permuting these experiences.

There are several interesting aspects of Mnih and colleagues' paper. First, the system performances are comparable to those of a human games tester. Second, the approach displays impressive adaptability. Although each system was trained using data from one game, the prior knowledge that went into the system design was essentially the same for all 49 games; the systems essentially differed only in the data they had been trained on. Finally, the main methods used have been around for several decades, making Mnih and colleagues' engineering feat all the more commendable.

What is responsible for the impressive performance of Mnih and colleagues' system, also reported for another DQN⁴? It may be largely down to improved function approximation using deep networks. Even though the size of the game screens produced by the emulator is reduced by the system to 84×84 pixels, the problem's dimensionality is much higher than that of most previous applications of reinforcement learning. Also, Q^* is highly nonlinear, which calls for a rich nonlinear function class to be used as an approximator. This type of approximation can be accurately made only using huge data sets (which the game emulator can produce), state-of-the-art function learning and considerable computing power.

Some fundamental issues remain open, however. Can we mathematically understand reinforcement learning from dependent data, and develop algorithms that provably work? Is it sufficient to learn statistical associations, or do we need to take into account the underlying causal structure, describing, say, which pixels causally influence others? This may help in finding relevant parts of the state space (for example, identifying which sets of pixels form a relevant entity, such as an alien in *Space Invaders*); in avoiding 'superstitious'

behaviour, in which statistical associations may be misinterpreted as causal; and in making systems more robust with respect to data-set shifts, such as changes in the behaviours or visual appearance of game characters^{3,5,6}. And how should we handle latent learning — the fact that biological systems also learn when no rewards are present? Could this help us to handle cases in which the dimensionality is even higher and the key quantities are hidden in a sea of irrelevant information?

In the early days of AI, beating a professional chess player was held by some to be the gold standard. This has now been achieved, and the target has shifted as we have grown to understand that other problems are much harder for computers, in particular problems involving high dimensionalities and noisy inputs. These are real-world problems, at which biological perception–action systems excel and

machine learning outperforms conventional engineering methods. Mnih and colleagues may have chosen the right tools for this job, and a set of video games may be a better model of the real world than chess, at least as far as AI is concerned. ■

Bernhard Schölkopf is at the Max Planck Institute for Intelligent Systems, 72076 Tübingen, Germany.
e-mail: bs@tuebingen.mpg.de

1. Mnih, V. et al. *Nature* **518**, 529–533 (2015).
2. Sutton R. S. & Barto A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
3. Watkins, C. J. C. H. *Learning from Delayed Rewards*. PhD thesis, Univ. Cambridge (1989).
4. Guo, X., Singh, S., Lee, H., Lewis, R. L. & Wang, X. *Adv. Neural Inf. Process. Syst.* **27** (2014).
5. Bareinboim, E. & Pearl, J. in *Proc. 25th AAAI Conf. on Artificial Intelligence* 100–108 (2011).
6. Schölkopf, B. et al. in *Proc. 29th Int. Conf. on Machine Learning* 1255–1262 (Omnipress, 2012).

BIODIVERSITY

The benefits of traditional knowledge

A study of two Balkan ethnic groups living in close proximity finds that traditional knowledge about local plant resources helps communities to cope with periods of famine, and can promote the conservation of biodiversity.

MANUEL PARDO-DE-SANTAYANA
& MANUEL J. MACÍÀ

Understanding how human groups obtain, manage and perceive their local resources — particularly the plants they use as food and medicine — is crucial for ensuring that those communities can continue to live and benefit from their local ecosystems in a sustainable way. The study of these complex interactions between plants and people is the aim of an integrative discipline known as ethnobotany, which is based on methods derived mainly from botany and anthropology¹. Most ethnobotanical research reveals that traditional knowledge about local edible and healing resources is suffering an alarming decline², especially in Europe³. However, writing in *Nature Plants*, Quave and Pieroni⁴ suggest that wild plants still have an essential role for communities living in the mountains of Kukës, one of the poorest districts of Albania. Their results also show how preserving local knowledge is linked to maintaining biodiversity.

The mountains of Kukës lie in the Balkans, a hotspot of cultural and biological diversity that has suffered major political and economic shifts over the past three decades. Quave and Pieroni studied two culturally

and linguistically distinct rural Islamic ethnic groups (the Gorani and Albanians) that, despite living in close proximity in this region and facing similar environmental and economic conditions, have remained relatively isolated from one another. The two groups use wild plants in different ways, giving the authors an opportunity to investigate the role of cultural factors in shaping how the local flora is understood and used in daily life, health practices and, ultimately, survival. Among the various quantitative techniques used, the authors designed a simple but innovative tool to compare the cultural similarities and differences between the two groups' use of plant species.

The researchers report significant variation in the plant species used for medicinal purposes by the two ethnic groups. A plausible explanation for this is that the spread of health-related lore requires a high degree of affinity, because trying a new remedy requires a great deal of trust⁵. Health is a sensitive topic, so people accept advice mainly from knowledgeable relatives or friends belonging to the same ethnic group⁶. Moreover, many traditional remedies have a highly symbolic component, and the mechanisms by which they are believed to bring about healing can lie — totally or partially — in the remedy's cultural meaning⁷.

Quave and Pieroni find only two species,



Figure 1 | The dog rose (*Rosa canina*) is used by both Gorani and Albanian ethnic groups.

Urtica dioica and *Rosa canina* (Fig. 1), that were widely used by both ethnic groups, and both species are edible. Generally, there was much more convergence in the food plants used by the two groups. The researchers suggest that this can be explained by the importance of wild edible species in ensuring food security. The robust local lore concerning these plants serves as a reservoir of knowledge, preparing the groups to cope with periods of famine or the scarcity of staple foods⁸. When food is scarce, cultural boundaries seem to be more permeable, because the survival of the group is at stake.

Another issue for consideration is the fact that some species are used medicinally by one group, but sold to plant traders by the other. People assign higher values to species that they use in their daily lives than to those that are harvested for marketing, which, as the authors point out, can have a major effect on the conservation of these resources. The group's relationship with the resource is much more intimate in the former case. Indeed, many regularly used plant species are of great cultural significance and have a prominent place in the local collective memory. They are part of local histories and narratives — they represent the essence, personality and identity of their community.

This study demonstrates that cultural values have a major effect on traditional local knowledge. Sustainable exploitation of local biodiversity is much more likely for resources that are emotionally valued than for those that are used in an impersonal way, as a source of income. A report published last year⁹ posited that many indigenous communities that have successfully conserved biodiversity in their locality do so by combining an extensive and experiential knowledge with an intensely respectful emotional engagement with nature. Furthermore, the report suggests that our

inclination to conserve biodiversity is a function of the number and intensity of our emotional attachments. Therefore, if traditional local knowledge is forgotten, biodiversity is also in danger of being lost, as is happening in some sacred forests and habitats that are in the process of being transformed and degraded¹⁰.

Studies such as Quave and Pieroni's can help to integrate traditional local knowledge with efforts to conserve biocultural diversity. By focusing on the point of view of people who are or have been deeply dependent on their local resources, these studies can promote culturally appropriate, sustainable development strategies. Unfortunately, this integration has received little support, and the implications of integration are yet to be properly evaluated. Future ethnobotanical studies should build on the innovative approach taken by Quave

and Pieroni when testing the role of cultural factors in the distribution and preservation of traditional local knowledge, by comparing the authors' results with larger data sets from informants and communities gathered in other regions worldwide.

Finally, the authors have demonstrated that quantitative techniques for analysing ethnobotanical data can lead to a deeper level of understanding within the discipline. We suggest that quantitative techniques should thus be further explored. Ethnobotany has a key part to play in studies of how ethnic groups can benefit from and coexist with their ecosystems. Policy and decision makers should take into account the views and traditions of local communities, particularly in rural regions with economic and social instability. ■

Manuel Pardo-de-Santayana and Manuel J. Macía are in the *Departamento de Biología (Botánica)*, Universidad Autónoma de Madrid, 28049 Madrid, Spain.
e-mail: manuel.pardo@uam.es

1. Schultes, R. E. & von Reis, S. *Ethnobotany: Evolution of a Discipline* (Chapman & Hall, 1995).
2. Reyes-García, V. *et al.* *Evol. Hum. Behav.* **34**, 249–257 (2013).
3. Pardo-de-Santayana, M., Pieroni, A. & Puri, R. K. (eds) *Ethnobotany in the New Europe: People, Health and Wild Plant Resources* (Berghahn, 2010).
4. Quave, C. L. & Pieroni, A. *Nature Plants* **1**, 14021 (2015).
5. Cámara-Leret, R., Paniagua-Zambrana, N., Svenning, J.-C., Balslev, H. & Macía, M. J. *J. Ethnopharmacol.* **158**, 58–65 (2014).
6. Menéndez-Baceta, G. *et al.* *J. Ethnopharmacol.* **161**, 116–127 (2015).
7. Moerman, D. E. & Jonas, W. B. *Ann. Intern. Med.* **136**, 471–476 (2002).
8. Johns, T. in *Eating on the Wild Side* (ed. Etkin, N. L.) 46–61 (Univ. Arizona Press, 2000).
9. Hunn, E. *Ethnobiol. Lett.* **5**, 146–150 (2014).
10. Verschuuren, B., Wild, R., McNeely, J. & Oviedo, G. (eds) *Sacred Natural Sites: Conserving Nature & Culture* (Earthscan, 2010).

STEM CELLS

Chasing blood

Many experiments have probed the mechanisms by which transplanted stem cells give rise to all the cell types of the blood, but it emerges that the process is different in unperturbed conditions. SEE LETTER P.542

SIDHARTHA GOYAL & PETER W. ZANDSTRA

Blood is one of the most dynamic tissues in the human body, with millions of cells being produced each second. But how blood-cell production occurs under unperturbed conditions, and which stem and progenitor cell types are responsible for stable maintenance of the blood, has been unclear. In this issue, Busch *et al.*¹ (page 542) use a genetic labelling strategy to gain insight into blood-cell

production under normal conditions. They reveal that an unexpected subset of blood stem cells is the major player in day-to-day blood-cell production.

Our understanding of haematopoiesis (the process through which all the cells of the blood are generated) is built on experiments in which blood cells are transplanted into recipients whose bone marrow (the source of blood-cell production) has been depleted through a process called myeloablation. The

cells' progeny are then tracked over time². Such work has helped to define key molecular and functional properties of blood stem cells, and to segregate the cells into different subsets, or compartments, on the basis of the time over which they contribute to haematopoiesis and on the lineages of cells that they can generate. At the top of the hierarchy are long-term haematopoietic stem cells (LT-HSCs), which give rise to short-term haematopoietic stem cells (ST-HSCs), both of which are named in accordance with the time over which they contribute to post-transplantation haematopoiesis. Both LT-HSCs and ST-HSCs exhibit self-renewal potential and typically give rise to all blood-cell lineages.

Only recently have we had the tools to extend our understanding of post-transplantation haematopoiesis to unperturbed conditions. Busch and colleagues' technique for studying unperturbed haematopoiesis is the cellular equivalent of classic 'pulse-chase' experiments, in which intracellular molecular components are traced by transient labelling of molecules to discover biosynthetic pathways. In the current study, the authors engineered mice such that LT-HSCs expressing the gene *Tie2* could be genetically labelled with a

fluorescent protein. Those cells and all their descendants will then fluoresce regardless of whether or not they express *Tie2*, allowing tracing of the cell lineages arising from the labelled stem cells. Following labelling, the authors performed a battery of assays at different times, and used mathematical modelling to analyse the resulting data (which were supported by *in vivo* stem-cell-transplantation experiments). This enabled them to define the rates of transition between blood-cell compartments, the length of time that cells spent in each compartment and the relative cell numbers in the different compartments (Fig. 1).

One of Busch and co-workers' central findings is that large numbers of ST-HSCs are responsible for most blood-cell production throughout the lifetime of the animal, with LT-HSCs participating to only a limited extent. This is perhaps not surprising, given that ST-HSCs proliferate faster than LT-HSCs. The authors also infer that at least 30% of LT-HSCs (around 5,000 cells in an adult mouse) go on to give rise to differentiated blood-cell lineages. This suggests that, in unperturbed conditions, the composition of the blood is highly polyclonal — that is, it is derived from many different stem cells.

The high level of polyclonality inferred by these experiments is in contrast to the findings of many transplant studies, in which a handful of individual LT-HSCs typically repopulate the entire blood system. However, to some extent, this dogma may be associated with the fact that often only a few transplanted cells successfully engraft in the bone marrow under normal experimental conditions. Busch *et al.* provide evidence suggesting that, unlike the case in unperturbed haematopoiesis, LT-HSCs are the more important contributors to the haematopoiesis that occurs during embryonic development or after treatment with a cytotoxic agent that kills blood cells. This finding suggests that feedback signalling from mature blood and progenitor cells³ to the LT-HSC compartment is key to controlling transition rates between compartments.

A paper published last year⁴ described the use of another genetic technique to track the contribution of a few thousand blood-stem-cell clones (each tagged with a different molecular signature) to the peripheral blood under normal conditions. A key finding of this work was that two blood-cell lineages, myeloid and lymphoid, were populated with cells harbouring different tags. This result agrees with Busch and colleagues' findings, suggesting that ST-HSCs that are beginning to become restricted to one lineage are the predominant source of haematopoiesis under unperturbed conditions. Also in agreement with Busch *et al.*, the paper reported that the peripheral blood was highly polyclonal.

Busch and colleagues' work, along with related papers^{4,5} examining unperturbed haematopoiesis *in vivo*, opens up several avenues for future research. A study⁶ involving transplantation of blood stem cells revealed a high degree of variability between the clonal populations in the post-transplantation blood, in both their size and their ability to give rise to different lineages. This variability can be attributed to various mechanisms, both intrinsic (different developmental potencies for the stem cells) and extrinsic (the surrounding microenvironment). Because of our limited understanding of clone sizes and lineage-contribution bias in unperturbed conditions, it is unclear how this post-transplantation variability relates to Busch and colleagues' observations under normal conditions. It could be that large numbers of contributing ST-HSCs are needed to ensure that robust polyclonality is maintained during normal haematopoiesis — something that is compromised when only limiting numbers of stem cells contribute to post-transplantation haematopoiesis. Alternatively, the reported post-transplantation clonal variability could be a stem-cell-intrinsic phenomenon specific to the post-transplantation environment. The ability to distinguish between LT- and ST-HSCs *in vivo* will be required to address this.

Another key area for investigation is

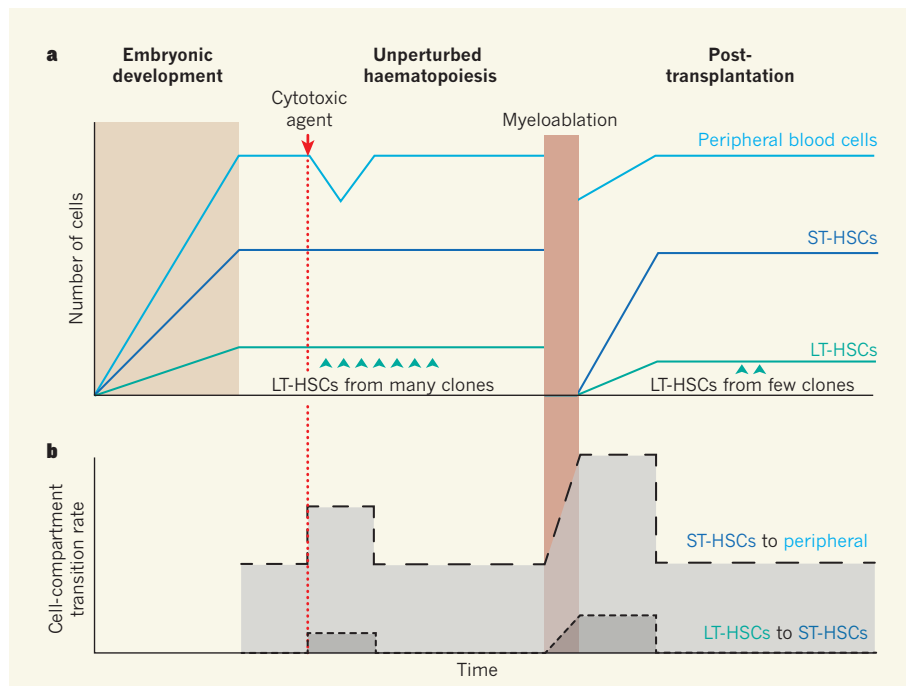


Figure 1 | A balance in the blood. During blood-cell production (haematopoiesis), long-term haematopoietic stem cells (LT-HSCs) give rise to short-term HSCs (ST-HSCs), which eventually produce the cells of the peripheral blood system. **a**, Busch *et al.*¹ investigated the numbers of each cell type over time. Numbers in each compartment rise during embryonic development to reach steady-state levels, which are maintained throughout adulthood. The death of peripheral blood cells owing to the introduction of a cytotoxic agent does not affect the numbers in either stem-cell compartment. If LT-HSCs and ST-HSCs are destroyed by myeloablation (depletion of bone-marrow cells), the numbers of peripheral blood cells dip, but quickly return to normal. LT-HSCs are derived from many clones under normal conditions, but fewer clones typically contribute to post-transplantation haematopoiesis. **b**, The authors also investigated the rate at which LT-HSCs transition to ST-HSCs, and at which ST-HSCs transition to peripheral blood cells in unperturbed and post-transplantation conditions (development was not studied). The rate of transition out of the ST-HSC compartment is consistently higher than the rate of transition from the LT-HSC compartment, implying that ST-HSCs are the major players in unperturbed haematopoiesis.



50 Years Ago

On August 20, 1964, one of us ... while trapping for small mammals near Listowel, County Kerry, caught an unusual 'mouse'. On subsequent examination it proved to be a member of the family Cricetidae, the bank vole, *Clethrionomys glareolus* Schreber — a family of mammals hitherto unknown from Ireland ... Comparisons of cranial characters were made with series of British mainland and continental specimens ... The only detectable difference is that the nasals are on average shorter and the condyle width greater than the British forms, but even here there is considerable overlap ... Investigations now being carried out are aimed at establishing the present distribution of this species in Ireland.

From *Nature* 27 February 1965

100 Years Ago

The Medical Committee of the British Science Guild has done a good work by its resolution condemning a notorious anti-vivisection advertisement. The object of the advertisement was to prevent our soldiers from being protected against typhoid fever. If it be asked why any one of the many anti-vivisection societies should behave in this way, we can only say with Dr. Watts that "Satan finds some mischief still for idle hands to do." ... Few of us are wanting to hear Pasteur called a charlatan; few of us are wanting anti-vivisection lectures and shops. Everybody is sure, who is capable of clear thinking, that our men of science are neither cruel nor stupid. But anti-vivisection cannot rest. It must find something to attack, something to abuse ... We hope that it will be many years before anti-vivisection emerges out of the public disgrace which it has brought upon itself.

From *Nature* 25 February 1915

why ST-HSCs can contribute to long-term haematopoiesis in unperturbed conditions, but contribute only transiently to post-transplantation haematopoiesis. Whether this reflects an effect of the post-transplant environment on these cells, or whether it is related to the increased proliferation in ST-HSCs during repopulation, is not clear. Finally, as Busch *et al.* mention, if these results extend to humans, efforts to capture the potential of ST-HSCs for clinical transplants could be valuable.

These areas of uncertainty cannot be addressed without the development of experimental techniques to mark prospective ST- and LT-HSCs *in vivo*, to quantitatively analyse the resulting lineages and model the data statistically⁷. Only then will we be able to fully interpret this complex and dynamic process. ■

Sidhartha Goyal is in the Department of Physics, University of Toronto, Toronto, Ontario M5S 1A7, Canada. **Peter W. Zandstra** is at the Institute of Biomaterials and Biomedical Engineering, University of Toronto, and at the Donnelly Centre for Cellular and Biomolecular Research, University of Toronto. e-mail: peter.zandstra@utoronto.ca

1. Busch, K. *et al.* *Nature* **518**, 542–546 (2015).
2. McCulloch, E. A., Siminovitch, L. & Till, J. E. *Science* **144**, 844–846 (1964).
3. Qiao, W. *et al.* *Mol. Syst. Biol.* **10**, 741 (2014).
4. Sun, J. *et al.* *Nature* **514**, 322–327 (2014).
5. Harrison, D. E., Lerner, C., Hoppe, P. C., Carlson, G. A. & Alling, D. *Blood* **69**, 773–777 (1987).
6. Dykstra, B. *et al.* *Cell Stem Cell* **1**, 218–229 (2007).
7. Blanpain, C. & Simons, B. D. *Nature Rev. Mol. Cell Biol.* **14**, 489–502 (2013).

This article was published online on 11 February 2015.

COSMOLOGY

A giant in the young Universe

Astronomers have discovered an extremely massive black hole from a time when the Universe was less than 900 million years old. The result provides insight into the growth of black holes and galaxies in the young Universe. [SEE LETTER P.512](#)

BRAM VENEMANS

It is commonly believed that every massive galaxy in the Universe harbours a super-massive black hole at its centre. These black holes are thought to have formed in the young Universe with initial masses of between 100 and 100,000 times the mass of the Sun¹. Over time, some of them have grown to be up to billions of solar masses by pulling (accreting) interstellar material from their surroundings and/or through merging with other black holes. The most massive black holes that have been found in the nearby Universe have masses of more than 10 billion solar masses^{2,3}. For comparison, our own Galaxy harbours a black hole with a mass of between 4 million and 5 million solar masses⁴. On page 512 of this issue, Wu *et al.*⁵ report the discovery of a supermassive black hole with a mass of a remarkable 12 billion solar masses, from a time when the Universe was only 875 million years old — that is, about 6% of its current age of 13.8 billion years.

Wu and colleagues identified this monster in optical and near-infrared imaging data because it was accreting gas at a high rate. This gas is pulled towards the black hole by gravity and can efficiently radiate away part of its potential energy. Accreting supermassive black holes can therefore be very bright, and can be seen across the Universe as luminous sources

termed quasars. Because the light coming from a very distant quasar takes billions of years to reach Earth, astronomers can observe such accreting black holes as they were when the Universe was young.

Theoretically, it is not implausible to find a black hole of more than 10 billion solar masses within 1 billion years after the Big Bang. But it is still surprising to uncover such a massive black hole in the early Universe. It must have been accreting gas at close to the maximum rate for most of its existence; the maximum rate is set by the pressure of the radiation emitted by the in-falling material. The prolonged period of almost maximum accretion is puzzling, because the strong radiation emitted by a quasar is generally assumed to be capable of halting accretion, limiting its existence to 10 million to 100 million years. The fact that the supermassive black hole has grown to 12 billion solar masses in less than a billion years implies that the radiation did not inhibit the high accretion.

In general, studies of supermassive black holes at the centres of nearby galaxies have revealed a tight correlation between the mass of the black hole and the total mass in stars of the galaxy hosting it⁶. Typically, the mass of a black hole is higher when it resides in a more massive galaxy, with the ratio of the black-hole mass to galaxy mass^{6,7} being about 0.14–0.5%.

Therefore, it has been suggested that the growth of both the black hole and the host galaxy are causally connected. If the relation between black-hole mass and host-galaxy mass were to hold true even in the distant Universe, we would expect the galaxy harbouring the 12-billion-solar-mass black hole to contain a whopping 4 trillion to 9 trillion solar masses in stars, which is the same as the most massive galaxies seen in the current Universe. Studying this host galaxy will give us a glimpse of how massive galaxies formed in the early Universe, and of the interplay between the formation of stars in the galaxy and the accretion onto its central black hole.

Intriguingly, the black hole discovered by Wu and collaborators is not only the most massive of its kind known in the early Universe, it is also, owing to the high accretion rate, by far the most luminous object detected at that cosmic epoch. The quasar can therefore be used as a means of learning about the distant cosmos. As the quasar's light travels towards observers on Earth, it passes through the gas of the intergalactic medium. This medium contains hydrogen, helium and various metals (elements heavier than helium that are produced inside stars), which leave an imprint on the spectrum of the quasar by absorbing a small amount of the quasar's light at specific wavelengths. The brighter the quasar, the more comprehensive the investigation of the intervening gas can be. Thus, the extreme brightness of the newly discovered quasar will allow the abundance of metals in the intergalactic medium of the early Universe to be measured in unprecedented detail. Such measurements will provide information about the star-formation processes at work shortly after the Big Bang, which produced these metals.

Finally, quasars as bright as the one reported here could easily be seen at larger distances from Earth than that of this quasar, and hence in an even younger Universe. Although accreting supermassive black holes become increasingly rare at earlier cosmic times⁸, current and future wide-field near-infrared imaging surveys should be able to uncover such objects. These giants of the Universe will provide the ideal targets from which to learn about the Universe during the first few hundred million years after the Big Bang. ■

Bram Venemans is at the Max Planck Institute for Astronomy, 69117 Heidelberg, Germany.
e-mail: venemans@mpia-hd.mpg.de

- Volonteri, M. *Astron. Astrophys. Rev.* **18**, 279–315 (2010).
- McConnell, N. J. *et al. Astrophys. J.* **756**, 179 (2012).
- van den Bosch, R. C. E. *et al. Nature* **491**, 729–731 (2012).
- Meyer, L. *et al. Science* **338**, 84–87 (2012).
- Wu, X.-B. *et al. Nature* **518**, 512–515 (2015).
- Kormendy, J. & Ho, L. C. *Annu. Rev. Astron. Astrophys.* **51**, 511–653 (2013).
- Håring, N. & Rix, H.-W. *Astrophys. J.* **604**, L89–L92 (2004).
- Willott, C. J. *et al. Astron. J.* **139**, 906–918 (2010).

QUANTUM PHYSICS

Teleportation for two

The ‘no-cloning’ theorem of quantum mechanics forbids the perfect copying of properties of photons or electrons. But quantum teleportation allows their flawless transfer — now even for two properties simultaneously. SEE LETTER P.516

WOLFGANG TITTEL

Suppose you see a beautiful table in a museum and you would like to have the same one at home. What could you do? One strategy is to accurately measure all its properties — its form (length, height and width) and its appearance (material and colour) — and then reproduce an identical copy for your living room. But this ‘measure-and-reproduce’ strategy would fail if the table were a quantum particle, such as a photon or an electron orbiting an atomic nucleus. The no-cloning theorem¹ of quantum mechanics tells us that it is impossible to copy such a particle perfectly. On page 516 of this issue, Wang *et al.*² show how to get around this apparent limitation of quantum physics. In a beautiful extension of previous experiments, they demonstrate how to transfer the values of two properties of a photon — the spin angular momentum

(the direction of the photon's electric field, generally referred to as polarization) and the orbital angular momentum (which depends on the field distribution) — through quantum teleportation onto another photon.

Quantum teleportation was proposed³ in 1993 and first demonstrated⁴ in 1997 for a single property of a photon (the polarization). It allows the flawless transfer of the unknown properties of an object onto a second object without contradicting the no-cloning theorem: the first object loses all its properties at the same time, that is, the properties are not ‘copied’ during quantum teleportation, they are transferred. However, the properties of the second object after this transfer remain unknown — all that is known is that they have been made identical to those of the first object before teleportation. What is more, the transfer does not happen instantaneously, a common mistake in the non-scientific literature.

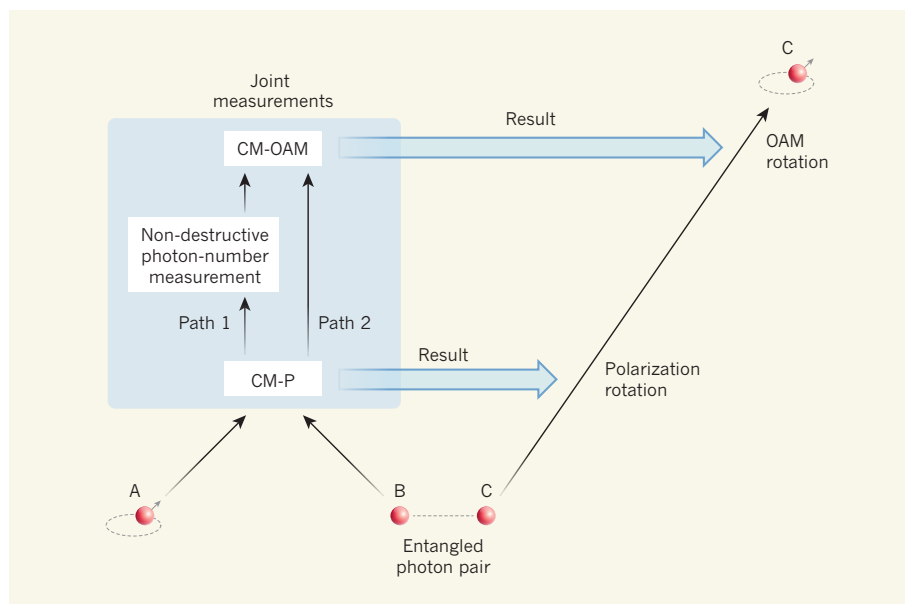


Figure 1 | Teleportation of photon polarization and orbital angular momentum. Photon A, whose polarization and orbital angular momentum are shown with a small arrow and an ellipse, respectively, is measured jointly with photon B, which is quantum-mechanically entangled with photon C. This act consists of: a comparative measurement of the polarizations of photons A and B (CM-P); a non-destructive verification that exactly one photon exits this measurement in path 1, and hence exactly one photon exits in path 2, given that two photons entered CM-P; and a comparative measurement of the orbital angular momenta of photons A and B (CM-OAM). The measurements result in the teleportation (that is, the transfer) of photon A's properties onto photon C. The transfer may require rotations of photon C's (unknown) polarization and orbital angular momentum, as determined by the outcomes of the comparative measurements. Wang *et al.*² have implemented all but the rotation steps in this transfer scheme. Teleporting the polarization alone does not require the non-destructive measurement, the CM-OAM, nor the rotation of photon C's orbital angular momentum.

In addition to the object (A) that carries the property to be teleported, quantum teleportation requires two more objects (B and C; Fig. 1). Objects B and C have to be entangled, which means that their properties are strongly correlated. For instance, the two photons B and C should have the same polarization, but the actual direction of their individual electric fields is not defined. Sounds weird? Just think, for example, that they are either both horizontally polarized or both vertically polarized, or both polarized at 45°. Photon A, whose polarization, say, will be teleported onto photon C, is measured jointly with photon B in a way that reveals, loosely speaking, the difference in the electric fields' directions without revealing the individual directions. What would we learn from getting, for instance, zero as the result? From the outcome of this comparative measurement, we know that the polarization of photon A equals that of photon B. Furthermore, from the entanglement of photons B and C, we know that the polarization of photon B equals that of photon C. Hence, we find that the electric field of photon C must now point in the same direction as that of photon A before the measurement.

Note that the outcome of the joint measurement could also have been different: for example, A and B are orthogonally polarized. Similar reasoning to that used before would lead to the conclusion that photon C's electric field is rotated by 90° with respect to that of photon A. Therefore, rotating it back would allow one to perfectly recover the original polarization encoded in photon A. In short, the joint measurement, possibly followed by a well-defined rotation of the (unknown) polarization of photon C, has allowed the teleporting (transferring) of the polarization property from photon A to photon C without error.

To demonstrate the teleportation of two properties, Wang and colleagues started with a single photon (photon A in Fig. 1) prepared in a combination of polarization and orbital angular momentum. Using high-intensity laser pulses that pass through a crystal, they also created a photon pair (photons B and C in Fig. 1) in a 'hyper-entangled' state, in which the photons are simultaneously entangled in the two properties to be teleported. Making two joint measurements (one per property) that compared the polarizations and the orbital angular momenta of photon A and photon B then led to the teleportation of photon A's properties onto photon C.

The biggest challenge for the researchers was the concatenation of the two joint measurements. It required, as an intermediate step, the verification that exactly one photon exited the first measurement (that of polarization) in each of the two possible paths leading to the second measurement (that of angular orbital momenta), without destroying the photons. The non-destructive detection of, say, a photon in path 1 can be implemented by teleporting

its orbital angular momentum onto another photon, which then enters the second joint measurement. This is because teleportation not only transfers a property from one photon to another, but also indicates that a photon existed. And, given that two photons entered (and hence left) the first comparative measurement, the non-destructive detection of a photon in path 1 also indicates that one photon was present in path 2 — exactly the requirement for the verification step. This step

This is an important step in understanding, and showcasing, one of the most profound and puzzling predictions of quantum physics.

needed another pair of photons (not shown in Fig. 1) entangled in their orbital angular momenta.

An interesting question is whether the demonstra-

ted method for the teleportation of two properties can be generalized to more properties. The authors affirm that this is possible in principle. However, the probability of the required joint measurements leading to a useful outcome becomes smaller and smaller as the number of properties (and thus of joint measurements) increases. Although the probability is half in the case of standard (single-property) teleportation, it is 1/32 for two properties, as shown for the first time by Wang and co-workers. Furthermore, it decreases to 1/4,096 when

teleporting an object that is described by three properties. Adding photons and photon detectors may increase the efficiency⁵, but this adds even more complexity to an already difficult measurement.

Even without these additional photons, the joint measurement becomes increasingly challenging as the number of properties increases: in the teleportation of two properties, a 'one-property teleporter' is used, and in the teleportation of three properties, a 'two-property teleporter' and a 'one-property teleporter' would be needed. You can guess what is required for the teleportation of N properties. Yet, Wang and colleagues' demonstration is an important step in understanding, and showcasing, one of the most profound and puzzling predictions of quantum physics. It may serve as a powerful building block for future quantum networks, which generally require teleportation units for the transmission of quantum data. ■

Wolfgang Tittel is in the Institute for Quantum Science and Technology, and the Department of Physics and Astronomy, University of Calgary, Calgary T2N 1N4, Canada.
e-mail: wtittel@ucalgary.ca

1. Wootters, W. K. & Zurek, W. H. *Nature* **299**, 802–803 (1982).
2. Wang, X.-L. *et al.* *Nature* **518**, 516–519 (2015).
3. Bennett, C. H. *et al.* *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
4. Bouwmeester, D. *et al.* *Nature* **390**, 575–579 (1997).
5. Grice, W. P. *Phys. Rev. A* **84**, 042331 (2011).

MOLECULAR BIOLOGY

RNA modification does a regulatory two-step

The m⁶A structural modification of RNA regulates gene expression. It has now been found to mediate an unusual control mechanism: by altering the structure of RNA, m⁶A allows a regulatory protein to bind to that RNA. SEE LETTER P.560

DOMINIK THELER & FRÉDÉRIC H.-T. ALLAIN

One of the most abundant modifications of messenger RNA is thought to be N⁶-methyladenosine (m⁶A), in which a methyl group is attached to the N6 position of adenine, an RNA base. The m⁶A modification has a role in regulating gene expression, and perturbations of this regulatory machinery are associated with human disease. But little is known about the mechanism by which the single methyl group of m⁶A exerts its effect. On page 560 of this issue, Liu *et al.*¹ report that m⁶A alters the secondary structure of RNA, allowing an RNA-binding protein to access the RNA sequence opposite the modification

and therefore to regulate expression.

Most experimental evidence for the mechanism and role of RNA modifications has been gathered in non-protein-coding RNAs. The m⁶A modification of mRNA was first described in 1974 (refs 2, 3), and subsequent studies quickly identified the methylase protein complex as the machinery that 'writes' m⁶A into mRNA (for reviews of m⁶A, see refs 4–6). Impairment of this complex leads to developmental arrest in several organisms.

After those early discoveries, not much was learnt about the role of m⁶A until the start of this decade, when m⁶A demethylase enzymes were identified as 'erasers' of this modification^{4–6}. These findings hinted at the dynamic

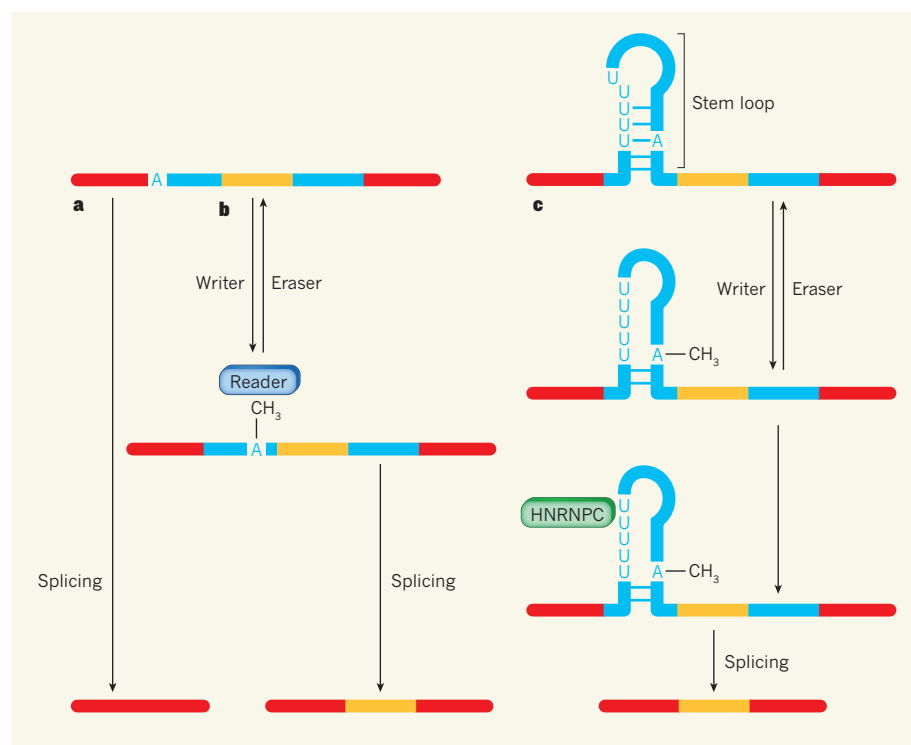


Figure 1 | Regulation of alternative splicing by the m⁶A RNA modification. Alternative splicing is the process by which different messenger RNAs are generated from transcripts of a single gene. Here, protein-coding sequences (exons) of an RNA transcript are shown in red and yellow, and non-coding sequences (introns) are in blue. **a**, In this example, splicing removes the introns and the exon trapped between them. **b**, When an adenine base (A) is methylated by a 'writer' protein to form the m⁶A modification (A-CH₃), direct binding of a 'reader' protein leads to splicing in which all the exons are included in the mRNA product. 'Eraser' proteins can reverse the formation of m⁶A. **c**, Liu *et al.*¹ report that, when m⁶A forms in a stem loop of RNA, the modification can alter the loop's structure by changing the base-pairing. The HNRNPC regulatory protein then binds to a stretch of uridine (U) nucleotides in the loop, also leading to inclusion of all the exons.

nature of m⁶A. Genetic alterations in the gene that encodes one of the erasers are associated with obesity and cancer in humans^{4–6}. In parallel, sequencing studies increased the number of known m⁶A sites from just a few to several thousand, and revealed that the sites are highly evolutionarily conserved in humans and mice^{4–6}.

'Reader' proteins of m⁶A have also been identified and found to be involved in gene expression. All reader proteins that have been biochemically validated to bind directly to the modification contain a hitherto poorly characterized RNA-binding domain called YTH. This domain binds to the methylated form of a given RNA with a several-fold higher affinity than it does for the unmethylated form, and the structural basis for this enhancement was reported last year⁷. Nevertheless, the mechanisms of action of these m⁶A readers are unknown.

The secondary structure of RNA provides another regulatory layer of gene expression, but is not strictly dependent on the primary sequence of bases. For example, RNA stem loops occur when two regions of a single-stranded RNA (ssRNA) molecule form a base-paired duplex (the stem) ending in an unpaired loop. Stem loops can act statically by attracting regulatory proteins to the stem or loop, or both. But they can also be dynamic, behaving

as RNA-based regulatory switches that influence cellular function by changing structure in response to different factors. These factors can range from physico-chemical parameters, including pH, temperature and the binding of ions and metabolites, to biological macromolecules such as nucleic acids and proteins⁸. Liu and colleagues report that there are several thousand such switches in mRNA, and that the switches are triggered by RNA modification.

Building on their discovery⁹ of an m⁶A site in the stem loop of a long non-coding RNA dubbed MALAT1, the authors found that the HNRNPC protein preferentially binds to the methylated form of this RNA; HNRNPC is an abundant ssRNA-binding protein involved in the regulation of several post-transcriptional gene-regulatory processes, such as alternative splicing. An HNRNPC-binding site¹⁰ consisting of a stretch of uridine (U) nucleotides is located in the stem loop opposite the m⁶A site. The researchers observed that methylation at the m⁶A site changes the structure of the stem loop by destabilizing base pairs and increasing the length of the single-stranded U-rich loop, thereby making the binding site more accessible to HNRNPC (Fig. 1).

Unusually, the m⁶A modification in MALAT1 does not recruit effector proteins

(readers) directly, but does so indirectly by altering the RNA's structure: methylation acts as a switch that alters the loop structure, which enables protein binding. This two-step process allows tight regulation of the events controlled by the m⁶A-switch. A study¹¹ published last month combined biophysics, structure determination and probing of the structural context of m⁶A *in vivo*. It revealed the propensity of m⁶A to alter secondary-structural elements formed from double-stranded RNA when located at the border between single-stranded and base-paired regions, as is observed in the case of MALAT1.

Liu *et al.* went on to characterize the abundance and functional roles of m⁶A-switches, identifying 2,798 of them with high confidence by observing which m⁶A sites display decreased HNRNPC binding when overall m⁶A levels in the transcriptome (the full set of transcribed RNAs) are reduced. They further characterized the interlinked actions of HNRNPC and m⁶A methylation by showing that the abundance of 5,251 different cellular RNA transcripts alters either when HNRNPC binding is inhibited or when m⁶A methylation is decreased. The authors also observed similar co-regulation by methylation and HNRNPC binding for many alternatively spliced exons (protein-coding regions of genes).

The findings are a huge advance in our understanding of how m⁶A controls gene expression, but several questions remain. Is HNRNPC the only RNA-binding protein to be recruited through such a mechanism upon m⁶A methylation, or are there others? Do the reader proteins involved in these m⁶A-switches further destabilize the secondary structure, or do they help to recruit the factors? Which cellular events trigger the switches? And, finally, do other RNA modifications in the vast transcriptome create RNA switches? We think that there are probably many more. ■

Dominik Theler and Frédéric H.-T. Allain
are at the Institute of Molecular Biology & Biophysics, Department of Biology, ETH Zurich, Zurich 8093, Switzerland.
e-mails: allain@mol.biol.ethz.ch;
dtheler@mol.biol.ethz.ch

1. Liu, N. *et al.* *Nature* **518**, 560–564 (2015).
2. Perry, R. P. & Kelley, D. E. *Cell* **1**, 37–42 (1974).
3. Desrosiers, R., Friderici, K. & Rottman, F. *Proc. Natl Acad. Sci. USA* **71**, 3971–3975 (1974).
4. Fu, Y., Dominissini, D., Rechavi, G. & He, C. *Nature Rev. Genet.* **15**, 293–306 (2014).
5. Meyer, K. D. & Jaffrey, S. R. *Nature Rev. Mol. Cell Biol.* **15**, 313–326 (2014).
6. Lee, M., Kim, B. & Kim, V. N. *Cell* **158**, 980–987 (2014).
7. Theler, D., Dominguez, C., Blatter, M., Boudet, J. & Allain, F. H.-T. *Nucleic Acids Res.* **42**, 13911–13919 (2014).
8. Dethoff, E. A., Chugh, J., Mustoe, A. M. & Al-Hashimi, H. M. *Nature* **482**, 322–330 (2012).
9. Liu, N. *et al.* *RNA* **19**, 1848–1856 (2013).
10. Cieniková, Z., Damberger, F. F., Hall, J., Allain, F. H.-T. & Maris, C. J. *Am. Chem. Soc.* **136**, 14536–14544 (2014).
11. Roost, C. *et al.* *J. Am. Chem. Soc.* **137**, 2107–2115 (2015).

Whole genomes redefine the mutational landscape of pancreatic cancer

Nicola Waddell^{1,2}, Marina Pajic^{3,4}, Ann-Marie Patch¹, David K. Chang^{3,5,6,7}, Karin S. Kassahn¹, Peter Bailey^{1,7}, Amber L. Johns³, David Miller¹, Katia Nones¹, Kelly Quek¹, Michael C. J. Quinn¹, Alan J. Robertson¹, Muhammad Z. H. Fadlullah¹, Tim J. C. Bruxner¹, Angelika N. Christ¹, Ivon Harliwong¹, Senel Idrisoglu¹, Suzanne Manning¹, Craig Nourse^{1,7}, Ehsan Nourbakhsh¹, Shivangi Wani¹, Peter J. Wilson¹, Emma Markham¹, Nicole Cloonan^{1,2}, Matthew J. Anderson¹, J. Lynn Fink¹, Oliver Holmes¹, Stephen H. Kazakoff¹, Conrad Leonard¹, Felicity Newell¹, Barsha Poudel¹, Sarah Song¹, Darrin Taylor¹, Nick Waddell¹, Scott Wood¹, Qinying Xu¹, Jianmin Wu³, Mark Pinese³, Mark J. Cowley³, Hong C. Lee³, Marc D. Jones^{3,7}, Adnan M. Nagrial³, Jeremy Humphris¹, Lorraine A. Chantrill³, Venessa Chin³, Angela M. Steinmann³, Amanda Mawson³, Emily S. Humphrey³, Emily K. Colvin³, Angela Chou^{3,8}, Christopher J. Scarlett^{3,9}, Andreia V. Pinho³, Marc Giry-Laterriere³, Ilse Rooman³, Jaswinder S. Samra^{10,11}, James G. Kench^{3,11,12}, Jessica A. Pettitt³, Neil D. Merrett^{5,13}, Christopher Toon³, Krishna Epari¹⁴, Nam Q. Nguyen¹⁵, Andrew Barbour¹⁶, Nikolajs Zeps^{17,18,19}, Nigel B. Jamieson^{7,20,21}, Janet S. Graham^{7,22}, Simone P. Niclou²³, Rolf Bjerkvig²⁴, Robert Grützmann²⁵, Daniela Aust²⁵, Ralph H. Hruban²⁶, Anirban Maitra²⁷, Christine A. Iacobuzio-Donahue²⁸, Christopher L. Wolfgang²⁹, Richard A. Morgan²⁶, Rita T. Lawlor^{30,31}, Vincenzo Corbo³⁰, Claudio Bassi³², Massimo Falconi^{32,33}, Giuseppe Zamboni^{31,33}, Giampaolo Tortora³⁴, Margaret A. Tempero³⁵, Australian Pancreatic Cancer Genome Initiative*, Anthony J. Gill^{3,11}, James R. Eshleman²⁶, Christian Pilarsky²⁵, Aldo Scarpa^{30,31}, Elizabeth A. Musgrove⁷, John V. Pearson^{1,2}, Andrew V. Biankin^{3,5,6,7}§ & Sean M. Grimmond^{1,7}§

Pancreatic cancer remains one of the most lethal of malignancies and a major health burden. We performed whole-genome sequencing and copy number variation (CNV) analysis of 100 pancreatic ductal adenocarcinomas (PDACs). Chromosomal rearrangements leading to gene disruption were prevalent, affecting genes known to be important in pancreatic cancer (*TP53*, *SMAD4*, *CDKN2A*, *ARID1A* and *ROBO2*) and new candidate drivers of pancreatic carcinogenesis (*KDM6A* and *PREX2*). Patterns of structural variation (variation in chromosomal structure) classified PDACs into 4 subtypes with potential clinical utility: the subtypes were termed stable, locally rearranged, scattered and unstable. A significant proportion harboured focal amplifications, many of which contained druggable oncogenes (*ERBB2*, *MET*, *FGFR1*, *CDK6*, *PIK3R3* and *PIK3CA*), but at low individual patient prevalence. Genomic instability co-segregated with inactivation of DNA maintenance genes (*BRCA1*, *BRCA2* or *PALB2*) and a mutational signature of DNA damage repair deficiency. Of 8 patients who received platinum therapy, 4 of 5 individuals with these measures of defective DNA maintenance responded.

Pancreatic cancer (PC) has a median survival of 6 months and a 5-year survival that remains less than 5% despite 50 years of research and therapeutic development¹. It is the fourth commonest cause of cancer death in Western societies and is projected to be the second leading cause within a decade. As a consequence, there is an urgent need to

better select patients for current therapies and develop novel therapeutic strategies.

Recent exome and CNV analyses of pancreatic ductal adenocarcinoma have revealed a complex mutational landscape^{2,3}. Activating mutations of *KRAS* are near ubiquitous and inactivation of *TP53*, *SMAD4*

¹Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Brisbane, Queensland 4072, Australia. ²QIMR Berghofer Medical Research Institute, Herston Road, Brisbane 4006, Australia. ³The Kinghorn Cancer Centre, Cancer Division, Garvan Institute of Medical Research, University of New South Wales, 384 Victoria St, Darlinghurst, Sydney, New South Wales 2010, Australia. ⁴St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, New South Wales 2010, Australia. ⁵Department of Surgery, Bankstown Hospital, Eldridge Road, Bankstown, Sydney, New South Wales 2200, Australia. ⁶South Western Sydney Clinical School, Faculty of Medicine, University of New South Wales, Liverpool, New South Wales 2170, Australia.

⁷Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, UK. ⁸Department of Anatomical Pathology, St Vincent's Hospital, Sydney, New South Wales 2010, Australia. ⁹School of Environmental & Life Sciences, University of Newcastle, Ourimbah, New South Wales 2258, Australia. ¹⁰Department of Surgery, Royal North Shore Hospital, St Leonards, Sydney, New South Wales 2065, Australia. ¹¹University of Sydney, Sydney, New South Wales 2006, Australia. ¹²Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Camperdown, New South Wales 2050, Australia. ¹³School of Medicine, University of Western Sydney, Penrith, New South Wales 2175, Australia. ¹⁴Department of Surgery, Fremantle Hospital, Alma Street, Fremantle, Western Australia 6160, Australia. ¹⁵Department of Gastroenterology, Royal Adelaide Hospital, North Terrace, Adelaide, South Australia 5000, Australia. ¹⁶Department of Surgery, Princess Alexandra Hospital, Ipswich Rd, Woollongabba, Queensland 4102, Australia. ¹⁷School of Surgery M507, University of Western Australia, 35 Stirling Highway, Nedlands 6009, Australia. ¹⁸St John of God Pathology, 12 Salvado Rd, Subiaco, Western Australia 6008, Australia. ¹⁹Bendat Family Comprehensive Cancer Centre, St John of God Subiaco Hospital, Subiaco, Western Australia 6008, Australia. ²⁰Academic Unit of Surgery, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow Royal Infirmary, Glasgow G4 0SF, UK. ²¹West of Scotland

Pancreatic Unit, Glasgow Royal Infirmary, Glasgow G31 2ER, UK. ²²Department of Medical Oncology, Beatson West of Scotland Cancer Centre, 1053 Great Western Road, Glasgow G12 0YN, UK. ²³Norlux Neuro-Oncology Laboratory, CRP-Santé Luxembourg, 84 Val Fleuri, L-1526, Luxembourg. ²⁴Norlux Neuro-Oncology, Department of Biomedicine, University of Bergen, Jonas Lies vei 91, N-5019 Bergen, Norway. ²⁵Departments of Surgery and Pathology, TU Dresden, Fetscherstr. 74, 01307 Dresden, Germany. ²⁶Department of Pathology, The Sol Goldman Pancreatic Cancer Research Center, the Johns Hopkins University School of Medicine, Baltimore, Maryland 21231, USA. ²⁷Departments of Pathology and Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston Texas 77030, USA. ²⁸The David M. Rubenstein Pancreatic Cancer Research Center and Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. ²⁹Department of Surgery, The Sol Goldman Pancreatic Cancer Research Center, the Johns Hopkins University School of Medicine, Baltimore, Maryland 21231, USA. ³⁰ARC-NET Centre for Applied Research on Cancer, University and Hospital Trust of Verona, Verona 37134, Italy. ³¹Department of Pathology and Diagnostics, University of Verona, Verona 37134, Italy. ³²Department of Surgery and Oncology, Pancreas Institute, University and Hospital Trust of Verona, Verona 37134, Italy. ³³Departments of Surgery and Pathology, Ospedale Sacro Cuore Don Calabria Negrar, Verona 37024, Italy. ³⁴Department of Oncology, University and Hospital Trust of Verona, Verona 37134, Italy. ³⁵Division of Hematology and Oncology, University of California, San Francisco, California 94122, USA.

*A list of participants and their affiliations appears at the end of the paper.

§These authors jointly supervised this work.

and *CDKN2A* occur at rates of >50%. The prevalence of recurrently mutated genes then drops to ~10% for a handful of genes involved in chromatin modification, DNA damage repair and other mechanisms known to be important in carcinogenesis; however, a long tail of infrequently mutated genes dominates, resulting in significant intertumoural heterogeneity. Faced with this diversity, it is not surprising that therapeutic development using an unselected approach to patient recruitment for clinical trials has been challenging^{2–4}.

Somatic structural rearrangement of chromosomes represents a common class of mutation that is capable of causing gene disruption (such as deletion or rearrangement), gene activation (for example, copy number gain or amplification) and the formation of novel oncogenic gene products (gene fusions). Many of these events actively drive carcinogenesis^{5,6} and in some instances present therapeutic targets. Early karyotyping⁷ and

more recent genomic sequencing of small numbers of primary tumours ($n = 3$) and metastases ($n = 10$) suggests that PDAC genomes contain widespread and complex patterns of chromosomal rearrangement^{8,9}.

Here we performed deep whole-genome sequencing of 100 PDACs and show that structural variation (variation in chromosomal structure) is an important mechanism of DNA damage in pancreatic carcinogenesis. We classify PDAC into four subtypes based on structural variation profiles and implicate molecular mechanisms underlying some of these events. Finally, as proof of concept, we use a combination of structural variation, mutational signatures and gene mutations to define putative biomarkers of therapeutic responsiveness for platinum-based chemotherapy, which are current therapeutic options for PDAC^{10–14}, and for therapeutics that target similar molecular mechanisms such as PARP inhibitors¹⁵ that are currently being tested in clinical trials.

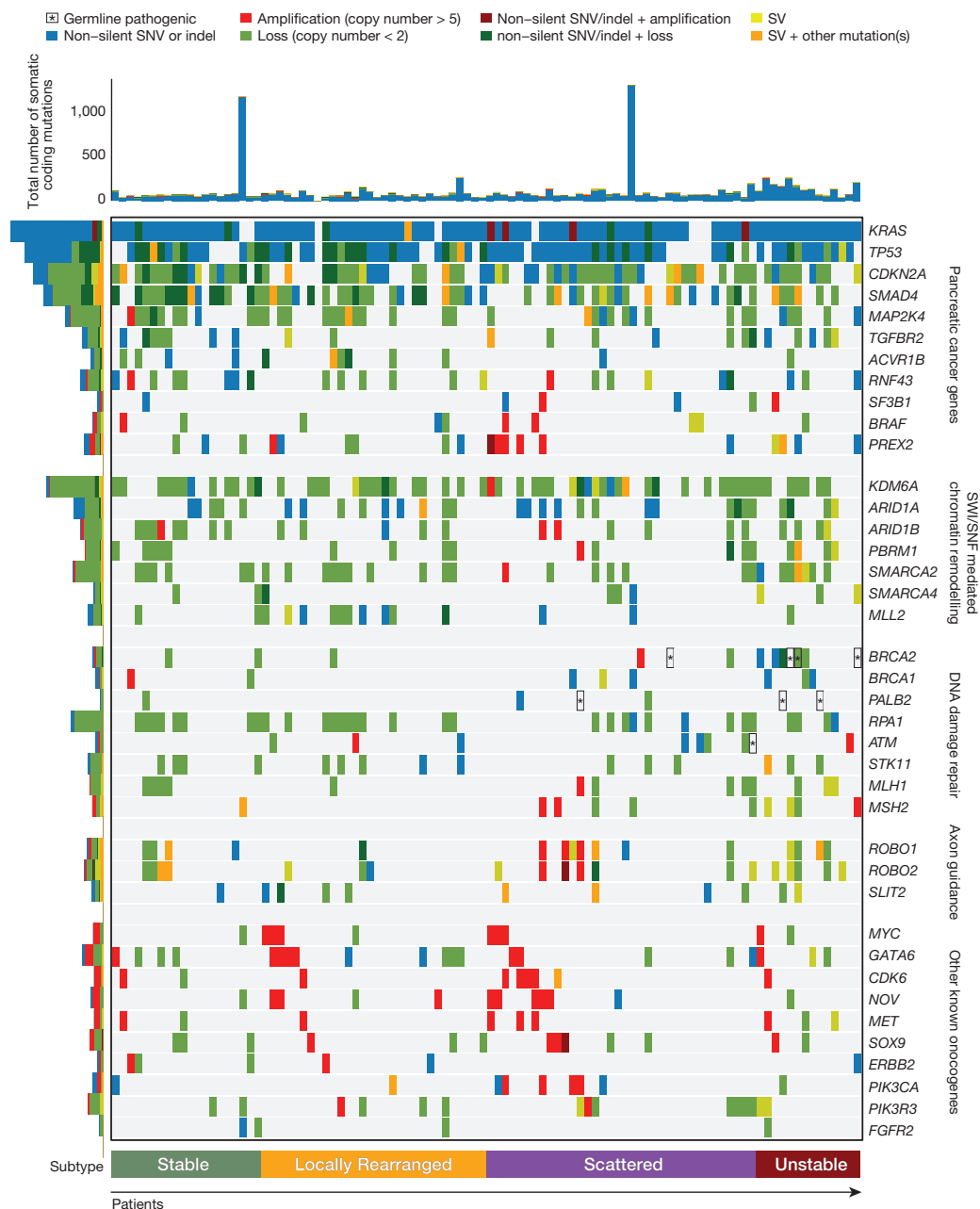


Figure 1 | Mutations in key genes and pathways in pancreatic cancer. The upper panel shows non-silent single nucleotide variants and small insertions or deletions. The central matrix shows: non-silent mutations (blue), copy number changes (amplification (>5 copies) represented in red and loss

represented in green) and genes affected by structural variants (SV, yellow). Pathogenic germline variants are highlighted with asterisk (*) symbols. The histogram on the left shows the number of each alteration in each gene.

Genomic landscape of pancreatic cancer

Patients were recruited and consent obtained for genomic sequencing through participating institutions of the Australian Pancreatic Cancer Genome Initiative (APGI; <http://www.pancreaticcancer.net.au>) as part of the International Cancer Genome Consortium (ICGC; <http://www.icgc.org>)¹⁶ (Supplementary Table 1). Array-based CNV was analysed using GAP¹⁷ and tumour cellularity estimated with qPure¹⁸. Whole-genome sequencing was performed on 100 primary PDACs with an epithelial cellularity of $\geq 40\%$ ($n = 75$), and complemented by cell lines derived from APGI participants ($n = 25$) to an average depth of $65\times$, and compared to the germline (average depth $38\times$) (Supplementary Table 2). Mutations were detected using qSNP¹⁹ and GATK and indels called with Pindel and GATK.

Point mutations and structural variation in PDAC

A total of 857,971 somatic point mutations and small insertions and deletions were detected in the cohort: 7,888 were non-silent mutations in 5,424 genes (Supplementary Tables 3 and 4). Orthogonal validation of $>3,000$ exonic mutations estimated the accuracy of mutation calls at $>95\%$ (Methods). Consistent with previous estimates²⁰, the average mutational burden across the cohort was 2.64 per Mb (range 0.65–28.2 per Mb). Somatic structural variants were identified with the qSV package, which uses multiple lines of evidence to define events (discordant pairs, soft clipping and split reads). Events verified using an orthogonal sequencing method were also included (Methods and Extended Data Fig. 1a). Where possible, these events were cross-referenced with CNV data (Methods). In total, 11,868 somatic structural variants were detected at an average of 119 per individual (range 15–558) (Supplementary Table 5 and Extended Data Fig. 1b). The majority of structural variants were intra-chromosomal (10,114) and were classified into 7 types: intra-chromosomal rearrangements (5,860), deletions (1,393), duplications (128), tandem duplications (179), inversions (1,629), fold-back inversions (579) and amplified inversions (346); inter-chromosomal translocations were less prevalent (1,754) (Supplementary Table 6). A total of 6,908 rearrangements directly disrupted gene sequences and

1,220 genes contained a breakpoint in 2 or more patients (Supplementary Table 7). Recurrent gene fusions were not detected: 1,236 structural variants led to the joining of two gene loci, however, only 183 of these events were fused in an orientation and frame that was capable of expressing a product, and none of these predicted fusion events occurred in more than one sample.

Genes affected by mutation and structural variation

Commonly mutated genes that characterize PDAC (*KRAS*, *TP53*, *SMAD4* and *CDKN2A*)^{2,3} were reaffirmed as significant using MutSig²¹ analysis (Supplementary Table 8). Combining structural variation events with deleterious point mutations increased the prevalence of inactivation events for *TP53* to 74% (3 structural variants and 71 mutations), 31% for *SMAD4* (9 structural variants and 22 mutations) and 35% for *CDKN2A* (11 structural variants and 24 mutations). Two additional genes not previously described in human PDAC (*KDM6A* and *PREX2*) had recurrent pathogenic mutations and structural variants at a rate of 10% or more. *KDM6A* is a SWI/SNF interacting partner that was identified in a pancreatic *sleeping-beauty* transposon mutagenesis screen²², and is mutated in RCC and medulloblastoma. In our cohort, *KDM6A* was inactivated in 18% of patients, (4 frame shifts, 1 in-frame deletion and 2 missense mutations, 5 structural variants and 8 homozygous deletions). In most cases ($n = 15$), both alleles of *KDM6A* were affected. The RAC1 guanine nucleotide exchange factor *PREX2*, mutated in melanoma²³ was inactivated in 10% of PDAC patients (1 frame shift, 1 splice site and 5 missense mutations, 2 structural variants and 1 homozygous deletion). In addition, the tumour suppressor gene *RNF43*, originally identified in cystic tumours of the pancreas, was inactivated in 10% of PDAC patients (4 frameshift and 4 nonsense mutations, 2 structural variants). Two of these PDACs had an associated intraductal papillary mucinous neoplasm (IPMNs). Recent studies have suggested that loss of functional *RNF43* may confer sensitivity to WNT inhibitors²⁴. Figure 1 shows the prevalence of aberrations in key driver genes and pathways in PDAC; implicating structural variation as an important mutational mechanism in pancreatic carcinogenesis.

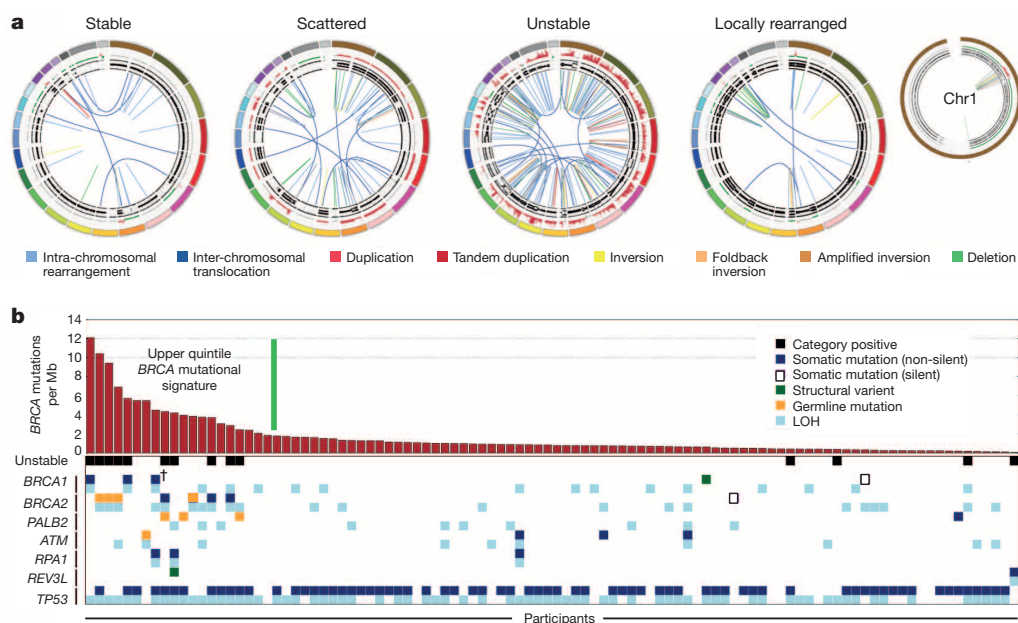


Figure 2 | Subtypes of pancreatic cancer. **a**, Subgroups of PDAC based on the frequency and distribution of structural rearrangements. Representative tumours of each group are shown. The coloured outer rings are chromosomes, the next ring depicts copy number (red represents gain and green represents loss), the next is the B allele frequency (proportion of the B allele to the total quantity of both alleles). The inner lines depict chromosome structural

rearrangements. **b**, The contribution of the *BRCA* mutational signature within each tumour ranked by prevalence (red bars). Unstable tumours are associated with a high *BRCA* mutation signature and deleterious mutations in *BRCA* pathway genes. The dagger (†) symbol indicates predicted only as possibly damaging by Polyphen2.

Table 10). Overlapping deleterious mutations in *BRCA1*, *BRCA2* and *PALB2* with unstable genomes and the *BRCA* mutational signature showed that mutations in these genes were associated with the top quintile of the *BRCA* mutational signature, and the majority (9 of 11) also exhibited unstable genomes (Fig. 3a).

Defective DNA repair without *BRCA* pathway mutations

Mutations in *BRCA* pathway genes accounted for approximately half of patients with a high *BRCA* mutational signature and/or an unstable genome (Fig. 3a). Hyper-methylation is known to play a role in silencing *BRCA1*, *BRCA2* and *PALB2* in some breast and ovarian cancers; however, high-density methylome array profiling of this cohort³⁰ allowed us to exclude this as a contributing mechanism. Single instances of biallelic, inactivating, somatic mutation was observed for two genes known to induce genomic instability and chemosensitivity when inactivated: *RPA1* (ref. 31) (splice site and loss of heterozygosity (LOH)), and the DNA polymerase zeta catalytic unit/*REV3L*³² (nonsense and LOH). We also detected mutations in other genes involved in DNA maintenance such as *ATM*, *FANCM*, *XRCC4* and *XRCC6* in tumours with an unstable genome or the *BRCA* mutational signature; however, they are yet to be causally linked to these genomic events or sensitivity to DNA-damaging agents.

Putative genotypes of platinum responsiveness

As the APCI was a prospective observational cohort study with extensive clinical follow-up, it was possible to track therapeutic responsiveness of participants that received chemotherapy when their disease recurred. At the time of analysis, 53 patients had documented recurrences and 25 received a variety of chemotherapeutic agents (Supplementary Table 11). This analysis was complemented through therapeutic testing of patient-derived xenografts (PDXs) generated from APCI participants. Overall, 8 patients received a platinum-based therapy and 7 PDXs were treated with gemcitabine and cisplatin (Fig. 3b). Of 5 patients with unstable genomes and/or a high *BRCA* mutational signature burden (designated

as ‘on-genotype’) 2 had exceptional responses (defined as complete radiological resolution of disease and normalization of CA19.9 levels³³), and 2 had robust partial responses based on *RECIST1.1* criteria³⁴ (Fig. 4a), while 3 patients who did not have any of these characteristics (‘off-genotype’) did not respond. These observations were supported by PDX studies where 2 of 3 on-genotype PDXs responded to cisplatin (one *BRCA2* mutant responded and one carrying bi-allelic inactivation of *RPA1*, which notably retained RAD51 foci (Extended Data Fig. 9) also responded. Another, with a mutational signature but not an unstable genome, and without a mutation in a *BRCA* pathway gene, did not respond. This compares to no responses in the 4 PDXs in the off-genotype group (Figs 3 and 4b). Combining patient and PDX response data, on-genotype tumours were associated with response to platinum-based therapy ($P = 0.0070$, Fisher’s exact test, Fig. 3b) (Supplementary Table 11).

Discussion

This study provides the most comprehensive description, to date, of the genomic events that characterize pancreatic cancer and demonstrates that structural variation is a prominent mechanism of genomic damage in this disease. It reinforces the importance of *KRAS*, *TP53*, *SMAD4*, *CDKN2A* and *ARID1A* gene mutations, in addition to numerous genes mutated at low prevalence. Recurrent mutations identified in *KDM6A* further highlights the role of chromatin modification and a broader role for aberrant WNT signalling is implicated through the relatively frequent inactivation of suppressor genes such as *ROBO1*, *ROBO2*, *SLIT2* and *RNF43*.

Structural variant analysis classifies PDAC into four subtypes with potential clinical relevance. A significant proportion of tumours contain amplifications and copy-number gains of known oncogenes, but most occur at low individual prevalence, suggesting significant diversity of mechanisms involved in PDAC progression. Several of these constitute known therapeutic targets with available inhibitors (*ERBB2*, *MET*, *FGFR1*). Others include: *GATA6*, which is known to be amplified in PDAC and correlates with poor survival in other cancer types³⁵; *PIK3CA*, which is

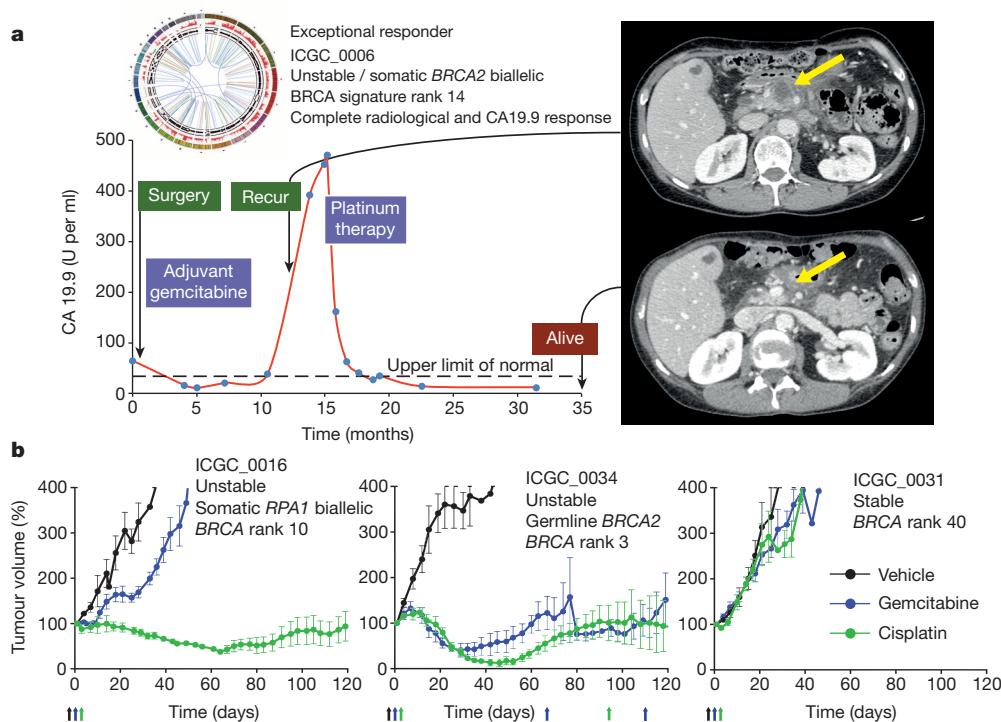


Figure 4 | Responses to platinum therapy. **a**, ICGC_0006: one of two patients who had an exceptional response to platinum-based therapy. Treatment of the recurrence with FOLFOX resulted in an exceptional response with recanalization of the portal vein which was previously obliterated by tumour and resolution of the mass with complete normalization of CA19.9

levels. **b**, Platinum responsiveness in patient-derived xenografts. Curves represent relative tumour volume ($0.5 \times \text{length} \times \text{width}^2$, y axis) over time (days, x axis). Arrows indicate drug treatment. Responses remain stable for >210 days. Error bars indicate standard error of the mean.

amplified in ovarian³⁶ and lung squamous cell carcinomas³⁷; *PIK3R3* amplified in ovarian cancer; and *CDK6*, amplified in oesophageal cancer. These may present opportunities for therapeutic intervention, either alone or in combination with other agents.

Multiple studies of platinum-based therapies in PDAC have shown borderline signals, and some meta-analyses show a benefit^{11,12}, suggesting that individual studies were underpowered, and that these signals could be driven by subgroups of responders. More recently, addition of oxaliplatin has shown efficacy in second line therapy¹⁴, and FOLFIRINOX, a platinum-containing combination therapy is emerging as a treatment option for advanced PDAC. Most patients do not receive this therapy due to its toxicity, or it is substantially modified³⁸. There are, however, significant responses in subgroups that are not well-defined^{39,40}, and improved survival reported in patients with germline *BRCA1* and *BRCA2* mutations who receive platinum-based therapies⁴¹. Defining biomarkers of platinum responsiveness would significantly alter current treatment approaches to PDAC and improve overall outcomes. Current patient recruitment strategies for clinical trials of PARP inhibitors, thought to target similar mechanisms, are mostly based on germline deleterious mutations of *BRCA1* and *BRCA2*. If we take into account mutations in *BRCA* pathway components, both germline and somatic, as well as putative surrogate measures of deficiencies in DNA maintenance, that is, unstable genomes and the *BRCA* mutational signature, germline mutations in *BRCA1* and *BRCA2* only account for as few as 4 of a potential 24 (17%), and only 4% of all patients. Genomic instability and *BRCA* mutational signature status based on whole-genome sequencing also provide independent evidence of putative deficiencies in DNA damage repair. It remains to be seen whether these surrogate measures are predictive of therapeutic response in the absence of *BRCA* or *PALB2* mutations. However, the presence of mutations in non-*BRCA* pathway genes that are associated with both genomic instability and chemosensitivity in 2/14 unstable tumours suggests that diagnostic whole-genome sequencing to detect surrogate measures of defects in DNA maintenance may ultimately be a better method of identifying potential responders to platinum and PARP inhibitor therapy.

The proof of concept data presented here suggest that mutations in *BRCA* pathway component genes and surrogate measures of defects in DNA maintenance (genomic instability and the *BRCA* mutational signature) have potential implications for therapeutic selection for pancreatic cancer. These data define a putative biomarker hypothesis that needs testing in a clinical trial, as these results are from a small number of patients selected based on high tumour cellularity; patients often received combination therapies, and the primary tumour was sequenced rather than the recurrence. As only selected gene sets can be tested in the clinic at this time, surrogate measures of molecular mechanisms identified using whole-genome sequencing can be used to inform individual gene selection for clinical use. As diagnostic genomic approaches continue to evolve and become more affordable, whole-genome sequencing may provide new opportunities in the clinic. However, there are significant hurdles still to overcome. These include the technical challenge of whole-genome sequencing using small diagnostic samples that are preserved in fixatives such as formalin, analytical demands and the return of results within a clinically relevant timeframe. Major initiatives are emerging that aim to address these challenges (such as Genomics England and the Scottish Genomes Partnership) to ultimately advance and assess these approaches for their potential to improve human health for many diseases including cancer.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 May; accepted 18 December 2014.

1. Vogelzang, N. J. *et al.* Clinical cancer advances 2011: annual report on progress against cancer from the American Society of Clinical Oncology. *J. Clin. Oncol.* **30**, 88–109 (2012).
2. Biankin, A. V. *et al.* Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* **491**, 399–405 (2012).

3. Jones, S. *et al.* Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**, 1801–1806 (2008).
4. Harada, T. *et al.* Genome-wide DNA copy number analysis in pancreatic cancer using high-density single nucleotide polymorphism arrays. *Oncogene* **27**, 1951–1960 (2008).
5. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).
6. Stephens, P. J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462**, 1005–1010 (2009).
7. Griffin, C. A. *et al.* Consistent chromosome abnormalities in adenocarcinoma of the pancreas. *Cancer Res.* **55**, 2394–2399 (1995).
8. Molenaar, J. J. *et al.* Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* **483**, 589–593 (2012).
9. Campbell, P. J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
10. Conroy, T. *et al.* FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N. Engl. J. Med.* **364**, 1817–1825 (2011).
11. Sultana, A. *et al.* Meta-analyses of chemotherapy for locally advanced and metastatic pancreatic cancer. *J. Clin. Oncol.* **25**, 2607–2615 (2007).
12. Ciliberto, D. *et al.* Role of gemcitabine-based combination therapy in the management of advanced pancreatic cancer: a meta-analysis of randomised trials. *Eur. J. Cancer* **49**, 593–603 (2013).
13. Heinemann, V., Boeck, S., Hinke, A., Labianca, R. & Louvet, C. Meta-analysis of randomized trials: evaluation of benefit from gemcitabine-based combination chemotherapy applied in advanced pancreatic cancer. *BMC Cancer* **8**, 82 (2008).
14. Oettle, H. *et al.* Second-line oxaliplatin, folinic acid, and fluorouracil versus folinic acid and fluorouracil alone for gemcitabine-refractory pancreatic cancer: outcomes from the CONKO-003 Trial. *J. Clin. Oncol.* **32**, 2423–2429 (2014).
15. Kaufman, B. *et al.* Olaparib monotherapy in patients with advanced cancer and a germline *BRCA1/2* mutation. *J. Clin. Oncol.* **33**, 244–250 (2015).
16. International Cancer Genome Consortium *et al.* International network of cancer genome projects. *Nature* **464**, 993–998 (2010).
17. Popova, T. *et al.* Genome alteration print (GAP): a tool to visualize and mine complex cancer genomic profiles obtained by SNP arrays. *Genome Biol.* **10**, R128 (2009).
18. Song, S. *et al.* qpure: a tool to estimate tumor cellularity from genome-wide single-nucleotide polymorphism profiles. *PLoS ONE* **7**, e45835 (2012).
19. Kassahn, K. S. *et al.* Somatic point mutation calling in low cellularity tumors. *PLoS ONE* **8**, e74380 (2013).
20. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
21. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
22. Mann, K. M. *et al.* *Sleeping Beauty* mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc. Natl Acad. Sci. USA* **109**, 5934–5941 (2012).
23. Berger, M. F. *et al.* Melanoma genome sequencing reveals frequent *PREX2* mutations. *Nature* **485**, 502–506 (2012).
24. Jiang, X. *et al.* Inactivating mutations of *RNF43* confer Wnt dependency in pancreatic ductal adenocarcinoma. *Proc. Natl Acad. Sci. USA* **110**, 12649–12654 (2013).
25. Korbel, J. O. & Campbell, P. J. Criteria for inference of chromothripsis in cancer genomes. *Cell* **152**, 1226–1236 (2013).
26. Tutt, A. *et al.* Absence of *Brca2* causes genome instability by chromosome breakage and loss associated with centrosome amplification. *Curr. Biol.* **9**, 1107–1110 (1999).
27. Jones, S. *et al.* Exomic sequencing identifies *PALB2* as a pancreatic cancer susceptibility gene. *Science* **324**, 217 (2009).
28. Hellebrand, H. *et al.* Germline mutations in the *PALB2* gene are population specific and occur with low frequencies in familial breast cancer. *Hum. Mutat.* **32**, E2176–E2188 (2011).
29. Nikkilä, J. *et al.* Heterozygous mutations in *PALB2* cause DNA replication and damage response defects. *Nature Commun.* **4**, 2578 (2013).
30. Nones, K. *et al.* Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of *SLIT-ROBO*, *ITGA2* and *MET* signaling. *Int. J. Cancer* **135**, 1110–1118 (2014).
31. Wang, Y. *et al.* Mutation in *Rpa1* results in defective DNA double-strand break repair, chromosomal instability and cancer in mice. *Nature Genet.* **37**, 750–755 (2005).
32. Doles, J. *et al.* Suppression of *Rev3*, the catalytic subunit of *Polε*, sensitizes drug-resistant lung tumors to chemotherapy. *Proc. Natl Acad. Sci. USA* **107**, 20786–20791 (2010).
33. Chang, D. K., Grimmer, S. M., Evans, T. R. J. & Biankin, A. V. Mining the genomes of exceptional responders. *Nature Rev. Cancer* **14**, 291–292 (2014).
34. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
35. Deng, N. *et al.* A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. *Gut* **61**, 673–684 (2012).
36. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
37. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
38. Peddi, P. F. *et al.* Multi-institutional experience with FOLFIRINOX in pancreatic adenocarcinoma. *JOP* **13**, 497–501 (2012).

39. Villarreal, M. C. *et al.* Personalizing cancer treatment in the age of global genomic analyses: *PALB2* gene mutations and the response to DNA damaging agents in pancreatic cancer. *Mol. Cancer Ther.* **10**, 3–8 (2011).
40. Showalter, S. L. *et al.* Identifying pancreatic cancer patients for targeted treatment: the challenges and limitations of the current selection process and vision for the future. *Expert Opin. Drug Deliv.* **7**, 273–284 (2010).
41. Golan, T. *et al.* Overall survival and clinical characteristics of pancreatic cancer in BRCA mutation carriers. *Br. J. Cancer* **111**, 1132–1138 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank C. Axford, M.-A. Brancato, S. Rowe, M. Thomas, S. Simpson and G. Hammond for central coordination of the Australian Pancreatic Cancer Genome Initiative, data management and quality control; M. Martyn-Smith, L. Braatvedt, H. Tang, V. Papangelis and M. Beilin for biospecimen acquisition; and D. Gwynne for support at the Queensland Centre for Medical Genomics. We also thank M. Hodgins, M. Debeljak and D. Trusty for technical assistance at Johns Hopkins University. N. Sperandio and D. Filippini for technical assistance at Verona University. We acknowledge the following funding support: National Health and Medical Research Council of Australia (NHMRC; 631701, 535903, 427601); Queensland Government (NIRAP); University of Queensland; Institute for Molecular Bioscience; Cancer Research UK (C596/A18076, C29717/A17263); Australian Government: Department of Innovation, Industry, Science and Research (DIISR); Australian Cancer Research Foundation (ACRF); Cancer Council NSW: (SRP06-01, SRP11-01, ICGC); Cancer Institute NSW: (10/ECF/2-26; 06/ECF/1-24; 09/ECF/2-40; 07/ECF/1-03; 10/CRF/1-01, 08/RSF/1-15, 07/ECF/1-28, 10/ECF/2-26, 10/FRL/2-03, 06/RSF/1-05, 09/RIG/1-02, 10/TPG/1-04, 11/REG/1-10, 11/ECF/3-26); Garvan Institute of Medical Research; Avner Nahmani Pancreatic Cancer Research Foundation; University of Glasgow; Cancer Research UK; Howat Foundation; R.T. Hall Trust; Petre Foundation; Philip Hemstitch Foundation; Gastroenterological Society of Australia (GESA); American Association for Cancer Research (AACR) Landon Foundation – INNOVATOR Award; Royal Australasian College of Surgeons (RACS); Royal Australasian College of Physicians (RACP); Royal College of Pathologists of Australasia (RCPA); Italian Ministry of Research (Cancer Genome Project FIRB RBAP10AHJB); Associazione Italiana Ricerca Cancro (12182); Fondazione Italiana Malattie Pancreas – Ministero Salute (CUP_J33G13000210001); Wilhelm Sander Stiftung 2009.039.2; National Institutes of Health grant P50 CA62924.

Author Contributions Biospecimens were collected at affiliated hospitals and processed at each biospecimen core resource centre. Data generation and analyses were performed by the Queensland Centre for Medical Genomics. Investigator contributions are as follows: A.V.B. and S.M.G. (concept and design); S.M.G., J.V.P. N.W., A.V.B. (project leaders); N.W., S.M.G., D.K.C., A.V.B. (writing team); J.V.P., S.M.G., N.W., A.L.J., P.B., S.S., K.S.K., N.K.W., P.J.W., A.M.P., F.N., B.P., E.M., O.H., J.L.F., C.L., D.T., S.W., Q.X., K.N., N.C., M.C.J.Q., M.J.A., M.Z.H.F., A.J.R., S.K., K.Q., M.Pi., H.C.L., M.J.C. and J.W. (bioinformatics); M.Pa., C.J.S., D.K.C., E.S.H., A.M.N., A.C., A.S., C.S., A.V.P., I.R., A.M.S., S.P.N., R.B. (preclinical testing); A.L.J., M.D.J., M.P., C.J.S., C.T., A.M.N., V.T.C., L.A.C., J.S.S., D.K.C., V.C., A.S., C.S., A.J.G., J.A.L., I.R., A.V.P., E.A.M. (sample processing and quality control); A.J.G., J.G.K., C.T., G.Z., A.S., D.A. R.H.H., A.M., C.A.I.-D., A.S. (pathology assessment); A.L.J., L.A.C., A.J.G., A.C., R.S.M., C.B., M.F., G.T., J.S.S., J.G.K., C.T., K.E., N.Q.N., N.Z., H.W., N.B.J., J.S.G., R.G., C.P., R.G., C.L.W., R.A.M., R.T.L., M.F., G.Z., G.T., M.A.T., A.P.G.J., J.R.E., R.H.H., A.M., C.A.I.-D., A.S. (sample collection and clinical annotation); D.M., T.J.C.B., A.N.C., I.H., S.I., S.M., C.N., E.N., S.W. (sequencing). All authors have read and approved the final manuscript.

Author Information BAM files and associated metadata in XML format have been uploaded to the European Genome-phenome Archive (EGA; <http://www.ebi.ac.uk/ega>) under accession number EGAS00001000154. All SNP array data is available via GEO (GSE61502). For more information about Australian Pancreatic Cancer Genome Initiative, see (<http://www.pancreaticcancer.net.au/apgi/collaborators>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.M.G. (sean.grimmond@glasgow.ac.uk) or A.V.B. (andrew.biankin@glasgow.ac.uk)

Australian Pancreatic Cancer Genome Initiative (Participants are arranged by institution.)

Garvan Institute of Medical Research Andrew V. Biankin^{1,2,22}, Amber L. Johns¹, Amanda Mawson¹, David K. Chang^{1,2,22}, Christopher J. Scarlett¹, Mary-Anne L. Brancato¹, Sarah J. Rowe¹, Skye H. Simpson¹, Mona Martyn-Smith¹, Michelle T.

Thomas¹, Lorraine A. Chantrill¹, Venessa T. Chin¹, Angela Chou¹, Mark J. Cowley¹, Jeremy L. Humphris¹, Marc D. Jones^{1,2}, R. Scott Mead¹, Adnan M. Nagrial¹, Marina Pajic¹, Jessica Pettit¹, Mark Pinese¹, Ilse Rooman¹, Jianmin Wu¹, Jiang Tao¹, Renee DiPietro¹, Clare Watson¹, Angela Steinmann¹, Hong Ching Lee¹, Rachel Wong¹, Andrea V. Pinho¹, Marc Giry-Laterriere¹, Roger J. Daly¹, Elizabeth A. Musgrove¹, Robert L. Sutherland^{1,†}; **Queensland Centre for Medical Genomics / Institute for Molecular Biosciences** Sean M. Grimmond³, Nicola Waddell³, Karin S. Kassahn³, David K. Miller³, Peter J. Wilson³, Ann-Marie Patch³, Sarah Song³, Ivon Harliwong³, Senel Idrisoglu³, Craig Nourse³, Ehsan Nourbakhsh³, Suzanne Manning³, Shivangi Wani³, Milena Gongora³, Matthew Anderson³, Oliver Holmes³, Conrad Leonard³, Darrin Taylor³, Scott Wood³, Christina Xu³, Katia Nones³, J. Lynn Fink³, Angelika Christ³, Tim Brunner³, Nicole Cloonan³, Felicity Newell³, John V. Pearson³, Peter Bailey³, Michael Quinn³, Shivashankar Nagaraj³, Stephen Kazakoff³, Nick Waddell³, Keerthana Krisnan³, Kelly Quek³, David Wood³, Muhammad Z. H. Fadlullah³; **Royal North Shore Hospital** Jaswinder S. Samra⁴, Anthony J. Gill⁴, Nick Pavlakis⁴, Alex Guminski⁴, Christopher Toon⁴; **Bankstown Hospital** Ray Asghari⁵, Neil D. Merrett⁵, Darren Pavey⁵, Amitabha Das⁵; **Liverpool Hospital** Peter H. Cosman⁶, Kasim Ismail⁶, Chelsie O'Connors⁶; **Westmead Hospital** Vincent W. Lam⁷, Duncan McLeod⁷, Henry C. Pleass⁷, Arthur Richardson⁷, Virginia James⁷; **Royal Prince Alfred Hospital** James G. Kench⁸, Caroline L. Cooper⁸, David Joseph⁸, Charbel Sandroussi⁸, Michael Crawford⁸, James Gallagher⁸; **Fremantle Hospital** Michael Texler⁹, Cindy Forest⁹, Andrew Laycock⁹, Krishna P. Epari⁹, Mo Ballal⁹, David R. Fletcher⁹, Sanjay Mukhedkar⁹; **Sir Charles Gairdner Hospital** Nigel A. Spry¹⁰, Bastiaan DeBoer¹⁰, Ming Chai¹⁰; **St John of God Healthcare** Nikolajs Zeps¹¹, Maria Beilin¹¹, Kynan Feeney¹¹; **Royal Adelaide Hospital** Nan Q. Nguyen¹², Andrew R. Ruzkiewicz¹², Chris Worthley¹², Chuan P. Tan¹², Tamara Debrincini¹²; **Flinders Medical Centre** John Chen¹³, Mark E. Brooke-Smith¹³, Virginia Papangelis¹³; **Greenslopes Private Hospital** Henry Tang¹⁴, Andrew P. Barbour¹⁴; **Envoi Pathology** Andrew D. Clouston¹⁵, Patrick Martin¹⁵; **Princess Alexandra Hospital** Thomas J. O'Rourke¹⁶, Amy Chiang¹⁶, Jonathan W. Fawcett¹⁶, Kellee Slater¹⁶, Shinn Yeung¹⁶, Michael Hatzifotis¹⁶, Peter Hodgkinson¹⁶; **Austin Hospital** Christopher Christophi¹⁷, Mehrdad Nikfarjam¹⁷, Angela Mountain¹⁷; **Victorian Cancer Biobank**¹⁸; **Johns Hopkins Medical Institutes** James R. Eshleman¹⁹, Ralph H. Hruban¹⁹, Anirban Maitra¹⁹, Christine A. Iacobuzio-Donahue¹⁹, Richard D. Schulick¹⁹, Christopher L. Wolfgang¹⁹, Richard A. Morgan¹⁹, Mary Hodglin¹⁹; **ARC-Net Centre for Applied Research on Cancer** Aldo Scarpa²⁰, Rita T. Lawlor²⁰, Stefania Beghelli²⁰, Vincenzo Corbo²⁰, Maria Scardoni²⁰, Claudio Bassi²⁰; **University of California, San Francisco** Margaret A. Tempero²¹; **University of Glasgow** Andrew V. Biankin^{1,2,22}, Sean M. Grimmond^{2,3}, David K. Chang^{1,2,22}, Elizabeth A. Musgrove², Marc D. Jones^{1,2}, Craig Nourse^{2,3}, Nigel B. Jamieson^{2,22}, Janet S. Graham^{2,22}; **Greater Glasgow & Clyde National Health Service** Andrew V. Biankin^{1,2,22}, David K. Chang^{1,2,22}, Nigel B. Jamieson^{2,22}, & Janet S. Graham^{2,22}, Karen Oien^{2,22} & Jane Hair^{2,22}

¹The Kinghorn Cancer Centre, Garvan Institute of Medical Research, 370 Victoria Street, Darlinghurst, Sydney, New South Wales 2010, Australia. ²Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Garscube Estate, Switchback Road, Bearsden, Glasgow, Scotland G61 1BD, UK. ³Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, University of Queensland, St Lucia, Queensland 4072, Australia. ⁴Royal North Shore Hospital, Westbourne Street, St Leonards, New South Wales 2065, Australia. ⁵Bankstown Hospital, Eldridge Road, Bankstown, New South Wales 2200, Australia. ⁶Liverpool Hospital, Elizabeth Street, Liverpool, New South Wales 2170, Australia. ⁷Westmead Hospital, Hawkesbury and Darcy Roads, Westmead, New South Wales 2145, Australia. ⁸Royal Prince Alfred Hospital, Missenden Road, Camperdown, New South Wales 2050, Australia. ⁹Fremantle Hospital, Alma Street, Fremantle, Western Australia 6959, Australia. ¹⁰Sir Charles Gairdner Hospital, Hospital Avenue, Nedlands, Western Australia 6009, Australia. ¹¹St John of God Healthcare, 12 Salvado Road, Subiaco, Western Australia 6008, Australia. ¹²Royal Adelaide Hospital, North Terrace, Adelaide, South Australia 5000, Australia. ¹³Flinders Medical Centre, Flinders Drive, Bedford Park, South Australia 5042, Australia. ¹⁴Greenslopes Private Hospital, Newdegate Street, Greenslopes, Queensland 4120, Australia. ¹⁵Envoi Pathology, 1/49 Butterfield Street, Herston, Queensland 4006, Australia. ¹⁶Princess Alexandra Hospital, 237 Ipswich Road, Woolloongabba, Queensland 4102, Australia. ¹⁷Austin Hospital, 145 Studley Road, Heidelberg, Victoria 3084, Australia. ¹⁸Victorian Cancer Biobank, 1 Rathdown Street, Carlton, Victoria 3053, Australia. ¹⁹Johns Hopkins Medical Institute, 600 North Wolfe Street, Baltimore, Maryland 21287, USA. ²⁰ARC-NET Center for Applied Research on Cancer, University of Verona, Via dell'Artiglierie, 19 37129 Verona, Province of Verona, Italy. ²¹University of California, San Francisco, 500 Parnassus Avenue, San Francisco, California 94122, USA. ²²Greater Glasgow and Clyde National Health Service, 1053 Great Western Road, Glasgow G12 0YN, UK.

†Deceased.

METHODS

Human research ethical approvals. Australian Pancreatic Cancer Genome Initiative: Sydney South West Area Health Service Human Research Ethics Committee, western zone (protocol number 2006/54); Sydney Local Health District Human Research Ethics Committee (X11-0220); Northern Sydney Central Coast Health Harbour Human Research Ethics Committee (0612-251M); Royal Adelaide Hospital Human Research Ethics Committee (091107a); Metro South Human Research Ethics Committee (09/QPAH/220); South Metropolitan Area Health Service Human Research Ethics Committee (09/324); Southern Adelaide Health Service/Flinders University Human Research Ethics Committee (167/10); Sydney West Area Health Service Human Research Ethics Committee (Westmead campus) (HREC2002/3/4.19); The University of Queensland Medical Research Ethics Committee (2009000745); Greenslopes Private Hospital Ethics Committee (09/34); North Shore Private Hospital Ethics Committee (09/324); Johns Hopkins Medical Institutions: Johns Hopkins Medicine Institutional Review Board (NA00026689). ARC-NET, University of Verona: approval number 1885 from the Integrated University Hospital Trust (AOUI) Ethics Committee (Comitato Etico Azienda Ospedaliera Universitaria Integrata) approved in their meeting of 17 November 2010 and documented by the ethics committee 52070/CE on 22 November 2010 and formalized by the Health Director of the AOUI on the order of the General Manager with protocol 52438 on 23 November 2010. Ethikkommission an der Technischen Universität Dresden (Approval numbers EK30412207 and EK357112012).

Animal experiment approvals. Mouse experiments were carried out in compliance with Australian laws on animal welfare. Mouse protocols were approved by the Garvan Institute/St Vincent's Hospital Animal Ethics Committee (ARA 09/19, 11/23 and 12/21 protocols). Female NOD/SCID/interleukin 2 receptor [IL2R] gamma (null) (NSG) mice and athymic Balb-c-nude mice were housed with a 12 h light, 12 h dark cycle, receiving food *ad libitum*.

Sample acquisition. Samples used were prospectively acquired and restricted to primary operable, non-pretreated pancreatic ductal adenocarcinoma. After ethical approval was granted, individual patients were recruited preoperatively and consented using an ICGC approved process. Immediately following surgical extirpation, a specialist pathologist analysed specimens macroscopically and samples of the tumour, normal pancreas and duodenal mucosa were snap frozen in liquid nitrogen (for full protocol see APGI website: <http://www.pancreaticcancer.net.au/>). The remaining resected specimen underwent routine histopathologic processing and examination. Once the diagnosis of pancreatic ductal adenocarcinoma was made, representative sections were reviewed independently by at least one other pathologist with specific expertise in pancreatic diseases (authors: A.G., D.M., R.H.H. and A.C.), and only those where there was no doubt as to the histopathological diagnosis were entered into the study. Co-existent intraductal papillary mucinous neoplasms in the residual specimen were not excluded provided the bulk of the tumour was invasive carcinoma, and the invasive carcinoma samples were used for sequencing. All samples were stored at -80°C . Duodenal mucosa or circulating lymphocytes were used for generation of germline DNA. A representative sample of duodenal mucosa was excised and processed in formalin to confirm non-neoplastic histology before processing. All participant information and biospecimens were logged and tracked using a purpose-built data and biospecimen information management system (Cansto Pancreas). Median survival was estimated using the Kaplan-Meier method and the difference was tested using the log-rank test. *P* values of less than 0.05 were considered statistically significant. Statistical analysis was performed using StatView 5.0 Software (Abacus Systems, Berkeley, CA, USA). Disease-specific survival was used as the primary endpoint.

Sample extraction. Samples were retrieved, and either had full face sectioning performed in OCT or the ends excised and processed in formalin to verify the presence of carcinoma in the sample to be sequenced and to estimate the percentage of malignant epithelial nuclei in the sample relative to stromal nuclei. Macrodissection was performed if required to excise areas of non-malignant tissue. Nucleic acids were then extracted using the Qiagen Allprep Kit in accordance with the manufacturer's instructions with purification of DNA and RNA from the same sample. DNA was quantified using Qubit HS DNA Assay (Invitrogen). Throughout the process, all samples were tracked using unique identifiers.

Patient material. One hundred matched normal and tumour derived samples were obtained from patients with PDAC. DNA was extracted from the samples using the Qiagen Allprep DNA/RNA mini kit method. Tumour cellularity was determined from SNP array data using qpure¹⁸. Clinical and sample data are summarized in (Supplementary Table 2). Patients were recruited and consent obtained for genomic sequencing through the Australian Pancreatic Cancer Genome Initiative (APGI) as part of the International Cancer Genome Consortium (ICGC)¹⁶.

Patient-derived cell line (PDCL) generation. The PDX-derived primary cell lines, named The Kinghorn Cancer Centre (TKCC) lines, were generated in the laboratory. All cell lines were profiled by short tandem repeat (STR) DNA profiling as unique (<http://www.cellbankaustralia.com>). Briefly, patient-derived tumours established

in immunocompromised mice were mechanically and enzymatically dissociated using collagenase (Stem Cell Technologies, USA) and plated onto flasks coated with 0.2mg ml^{-1} rat tail collagen (BD Biosciences, USA). Subsequently, epithelial cultures were enriched and purified using a FACS Aria III Cell sorter (BD Biosciences, USA), using a biotinylated anti-mouse MHC1 antibody (1:200 dilution; eBiosciences, USA) coupled with Streptavidin AlexaFluor 647 secondary step (1:1,000; Invitrogen, USA) and anti-mouse CD140a-PE antibody (1:300; BD Biosciences, USA) to remove mouse stroma. Dead cells were removed using propidium iodide (Sigma-Aldrich, Australia). Following establishment, all patient-derived (TKCC) cell lines were profiled by short tandem repeat (STR) DNA profiling as unique (<http://www.cellbankaustralia.com>). **Sequencing.** DNA ($1\text{ }\mu\text{g}$) was diluted to $52.5\text{ }\mu\text{l}$ in DNase-/RNase-free molecular biology grade water before fragmentation to approximately 300 bp using the Covaris S2 sonicator with the following settings Duty Cycle 10%, intensity 5, cycles per burst 200, time 50 s or 45 s for PCR-Free libraries. Following fragmentation libraries for sequencing were prepared using the standard Illumina library preparation technique of end-repair, adenylate 3' ends, indexed adaptor ligation, size selection and finally PCR enrichment for adaptor ligated library molecules following the manufacturer's recommendations (Part no. 15026486 Rev. C July 2012). A subset of libraries was generated omitting the final PCR enrichment step to generate PCR-Free libraries as per the manufacturer's recommendations (Part no. 15036187 Rev. A Jan 2013). For standard libraries commercially available TruSeq DNA LT Sample Prep Kit v2 (Catalogue no. FC-121-2001) were used for all steps with the following exceptions. Size selections of the Adaptor Ligated fragments were completed using two rounds of SPRI bead purifications (AxyPrepMag PCR Clean-up Catalog no. MAG-PCR-CL-250) using a final bead to DNA volume ratio of 0.60:1 followed by 0.70:1, selecting for molecules with an average size of 500 bp. Size-selected libraries were then amplified for a total of 8 cycles of PCR to enrich for DNA fragments both compatible with sequencing and containing the ligated indexed adaptor. For PCR-Free libraries commercially available TruSeq PCR-Free DNA LT Sample Preparation Kit (Catalog no. FC-121-3001 and FC-121-3002) was used following the 350 bp library LT protocol for all steps with no modifications. The final whole-genome libraries were qualified (amplified and PCR-Free libraries) and quantified (amplified libraries only) via the Agilent BioAnalyser 2100 (Catalog ID: G2940CA) instrument using the DNA High Sensitivity kit (Catalog ID: 5067-4626). Quantification of PCR-Free libraries was performed using the KAPA Library Quantification Kits For Illumina sequencing platforms (Kit code KK4824) in combination with Life Technologies Viia 7 real time PCR instrument.

Whole genome libraries were prepared for cluster generation by cBot (catalogue no. SY-301-2002) and sequencing as per the manufacturer's guidelines. Individual libraries were clustered on a single lane of a HiSeq v3 flowcell using the TruSeq PE Cluster Kit v3-cBot-HS kit (Catalogue no. PE-401-3001). Illumina supplied control library PhiX (10 pM) was spiked into each lane at a concentration of 0.3% to provide real time analysis metrics. Final library concentrations of 8 pM (amplified) and 14 pM (PCR-free) were used for cluster generation. Clustered flowcells were sequenced on the Illumina HiSeq 2000 instrument (HiSeq control software v1.5/Real Time Analysis 1.13) using TruSeq SBS Kit v3-HS (200 cycles, Catalog no. FC-401-3001). Paired reads each of 101 bp were generated for all libraries and in total approximately 220-million paired reads were generated per lane, in line with the manufacturer's specification. Real time analysis of the control library PhiX showed cluster density, error rates, quality scores, mapping rates and phasing rates were also in line with published specifications.

Sequence alignment and data management. Sequence data was mapped to a genome based on the Genome Reference Consortium (<http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/>) GRCh37 assembly using BWA⁴². Multiple BAM files from the same sequence library were merged and within library duplicates were marked. Resulting final BAMs were used as input into variant calling. All BAM files have been deposited in the EGA (Accession number: EGAS 00001000154).

Copy number analysis. Matched tumour and normal patient DNA was assayed using Illumina SNP BeadChips as per manufacturer's instructions (Illumina, San Diego CA) (HumanOmni1-Quad or HumanOmni2.5-8 BeadChips). SNP arrays were scanned and data was processed using the Genotyping module (v1.8.4) in Genomestudio v2010.3 (Illumina, San Diego CA) to calculate B-allele frequencies (BAF) and logR values. GenoCN⁴³ and GAP¹⁷ were used to call somatic regions of copy number change – gain, loss or copy neutral LOH. Recurrent regions of copy number change were determined and genes within these regions were extracted using ENSEMBL v70 annotations.

Identification of structural variations. Somatic structural variants were identified using the qSV tool (manuscript in preparation). qSV uses independent lines of evidence to call structural variants including discordant reads, soft clipping and split read. Breakpoints are also identified using both *de novo* assembly of abnormally mapping reads and split contig alignment to enhance break point resolution. Depending on the level of evidence qSV bins calls into different categories and calls were

considered high confidence if: (i) they were category 1 and therefore contain multiple lines of evidence (discordant pairs, soft clipping on both sides and split reads); (ii) they were category 2 and therefore there was 2 lines of evidence: discordant pairs (both breakpoints) and soft clipping; or discordant pairs (both breakpoints) and split read; or soft clipping (double sided) and split read; (iii) they were category 3 with 10 or more supporting events (discordant read pairs or soft clipping at both ends). Only high confidence calls were used in further downstream analysis. Copy number variation was estimated using SNP arrays and the GAP tool¹⁷. Depending on the read pair types supporting an aberration or the associated of copy number events each structural variant was classified as: deletion, duplication, tandem duplication, foldback inversion, amplified inversion, inversion, intrachromosomal or translocation. Essentially, the type of rearrangement is initially inferred from the orientation information of discordant read pairs, soft clipping clusters and assembled contigs which span the breakpoints. This allows identification of 4 groups of events: duplications/intra-chromosomal rearrangements, deletions/intra-chromosomal rearrangements, inversions and inter-chromosomal translocations. Boundaries of segments of copy number that occur in close proximity to each breakpoint were then used to aid further classification of the events. Structural variants with breakpoints that flanked a copy number segment of loss were annotated as deletions. Duplications and inversions associated with increases in copy number enabled the characterization of tandem duplications and amplified or foldback inversions. Events within the same chromosome which linked the ends of copy number segments of similar copy number levels were often identified and were called intra-chromosomal rearrangements.

Events were then annotated if they were within 100 kb of a centromere or telomere and genes which were affected by breakpoints were annotated using ENSEMBL v70. Structural variants and copy number data were visualized using *circos*⁴⁴.

The landscape of structural rearrangements in pancreatic ductal adenocarcinoma. In total 11,868 structural variants were detected within the 100 PDAC cohort with an average of 119 events per patient (range 15–558). Each event was classified into one of 8 categories: deletion, duplication, tandem duplication, foldback inversion, amplified inversion, inversion, intra chromosomal and translocation. Within the cohort there was inter patient heterogeneity in terms of total number of events (range of events per patient 15–558) and proportion of event type (Extended Data Fig. 1).

Classification of subtypes based on the pattern or structural rearrangements. Each tumour was classified into one of four subtypes based on the volume of events, the predominance of specific types of structural rearrangement events and the distribution of events across the genome in each patient. In addition to counting structural variation events, two analyses were carried out to detect localized events. Non-random chromosomal clustering of structural variants was detected using an approach originally described by Korbel and Campbell²⁵. Significant clustering of structural variation events was determined by a goodness-of-fit test against the expected exponential distribution of (with a significance threshold of < 0.0001). Highly focal events were detected using an adaptation of a method⁴⁵ where chromosomes with a high structural variant mutation rate per Mb exceeded 5 times the length of the interquartile range from the 75th percentile of the chromosome counts for each patient. The rules used to determine these subtypes are as follows:

Stable These tumours contain few structural rearrangements (< 50) which are located randomly through the genome.

Locally rearranged The intra-chromosomal rearrangements in these tumours are not randomly positioned through the genome, instead they are clustered on one or few chromosomes. To correct for the different chromosome lengths, the number of events per Mb was calculated for each chromosome within each tumour. Tumours were considered locally rearranged if they harboured at least 50 somatic events within the genome and contained a locally rearranged chromosome. Chromosomes were considered locally rearranged if the number of intrachromosomal events exceeded 5 times the length of the interquartile range from the 75th percentile of the chromosome counts per Mb for that patient. The events in the locally rearranged tumours are broadly comprised of either: (1) focal amplifications—the majority of events are gain (tandem duplication, duplication, foldback inversion or amplified inversion) or (2) complex rearrangements—the events are part of a complex event such as chromothripsis or breakage–fusion–bridge.

Scattered These tumours contain 50–200 structural rearrangements which are scattered throughout the genome.

Unstable These tumours are massively rearranged as they contain > 200 structural rearrangements which are generally scattered throughout the genome.

Classification of complex localized events. Evidence of clustering of breakpoints was estimated as proposed by Korbel and Campbell²⁵. Chromosomes with clustering of structural variants were reviewed for evidence of chromothripsis (oscillation of copy number, random joins and retention of heterozygosity) and breakage–fusion–bridge (BFB for loss of telomeric region with neighbouring highly amplified region with inversions).

Verification of structural variations. We used two methods of verification for structural variants: (1) an *in silico* approach, which considers events with multiple lines of evidence (qSV category 1: discordant pairs, soft clipping on both sides and split read evidence) as verified, as well as events which were associated with a copy number change (gain or loss) and (2) orthogonal sequencing methods including SOLiD long mate pair and capillary sequencing.

Long mate pair sequencing and verification of structural rearrangements. Long mate-paired libraries were made according to Applied Biosystems Mate-Paired Library Preparation 5500 Series SOLiD systems protocol using 5 µg of DNA which was sheared using the CovarisS220 System. Long mate pair libraries were sequenced using the SOLiD v4 (Applied Biosystems). Sequence data was mapped to a genome based on the Genome Reference Consortium (<http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/>). GRCh37 assembly using *bioScope* v1.2.1 (Applied Biosystems). Each sample was sequenced to an average non-redundant physical coverage of $180\times$ (64–333) in the tumour and $187\times$ (52–503) in the control sample. Structural rearrangements were determined by analysing clusters of discordant read pairs using the qSV tool. Events identified by Hiseq sequencing were considered verified if the right and left breakpoint of these events were within 500 bases of the right and left breakpoint of an event identified by SOLiD sequencing.

PCR and capillary sequencing for verification of structural rearrangements. For PCR and capillary sequencing PCR primers were designed with primer BLAST (NCBI) to span the predicted breakpoint, primers were designed with primer BLAST (NCBI). PCR was carried out in the tumour and matched normal genomic DNA using, respectively, a 25 or 50 µl reaction volume composed of 22 or 44 µl of Platinum Taq DNA polymerase (Invitrogen, Carlsbad, Ca), 2 or 4 µl of 10 µM primer (Integrated DNA Technology) and 1 or 2 µl of genomic DNA as template (1 ng µl⁻¹). The following parameters were used for the PCR. Initial denaturation at 94 °C for 2 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 60 °C for 30 s and extension at 68 °C for 1 min; followed by final extension at 68 °C for 15 min. PCR products were visualized by gel electrophoresis and classified into one of four categories: (1) validated—strong and specific PCR band of the expected size was observed only in the tumour and not in the normal sample, this indicates a somatic rearrangement; (2) germline—clear PCR band of the expected size both in the tumour and normal; (3) not validated—PCR yields smears or multiple bands, this potentially indicates non-specific primer pair; (4) not tested—no PCR band was observed in tumour and normal.

Verification of structural variations—results. In total 7,105 events were verified *in silico*. Of these 5,666 events contained multiple lines of evidence (qSV category 1), 2,904 events were associated with a copy number change (events classified as deletion, duplication, tandem duplication, amplified inversion and foldback inversion) and 1,871 contained multiple lines of evidence and were associated with a copy number change.

We also verified structural variant events using long mate pair resequencing (SOLiD paired 50 bp) or sequencing of a different sample from the same patient of 33 tumours. Using this approach 1,924 events were confirmed and the verification status of structural variant events was recorded in Supplementary Table 5 in the “validation_status_id” column where 0 = untested and 1 = verified. In total 7,228 of the 11,868 events identified (61%) were verified (Supplementary Table 5 and Extended Data Fig. 1) the remaining events remain untested.

Identification of substitutions and small insertion/deletions. Substitutions are called using 2 variant callers: qSNP¹⁹ an in-house heuristics-driven somatic/germline caller; and GATK⁴⁶ which is a Bayesian caller. The two callers were chosen because they use very different calling strategies and while each maybe subject to artefacts (as are all variant callers), they will be subject to different artefacts. Each compared variant falls into one of three categories: seen only by qSNP, seen only by GATK, and seen by both qSNP and GATK. Mutations identified by both callers or those that were unique to a caller and verified by an orthogonal sequencing approach were considered high confidence and used in all subsequent analyses (Supplementary Table 3). Small indels (< 200 bp) were identified using Pindel⁴⁷; each indel was visually inspected in the Integrative Genome Browser (IGV)⁴⁸. Once somatic mutations were called, their effects on any alternative transcripts were annotated using a local install of the Ensembl database (v70) and the Ensembl Perl API.

Verification of substitutions and small insertion/deletions. In total 3,304 of the 10,335 events identified were verified (Supplementary Tables 3 and 12) the remaining events remain untested. Substitutions and indels were verified using orthogonal sequence data which included data produced on different sequencing platforms (Hiseq or SOLiD exome or long mate pair SOLiD sequencing) or data from related nucleotide samples (RNA-seq). For example, if orthogonal tumour sequence data was available (DNA from a cell line, RNA from the primary sample etc.) and a somatic variant was also observed in the second tumour sample then that would add support for the variant. It should be noted that tumour samples can only be used to support an existing somatic variant and the absence of a called variant in a second tumour sample does not discredit the original call. Conversely, a second normal sample will

only discredit somatic variants and the absence of the called variant in the second normal does not support the original call. This approach is designed to be conservative. In order to be considered for verification, an additional BAM should have a minimum of 10 reads at the variant position, and at least two reads must show the variant. If multiple additional BAMs are available, each BAM votes independently and the concordance of the votes is used to classify the verification of the variant. Each variant examined by qVerify is assigned to one of four categories:

(1) Verified—one or more additional tumour BAMs showed evidence of the variant and no additional normal BAMs showed the variant.

(2) False Positive—one or more additional normal BAMs showed evidence of the variant indication that it is likely to be a germline variant.

(3) Mixed—across multiple additional BAMs, there was conflicting evidence—one or more additional tumour BAMs showed the variant as did one or more additional normal BAMs. This could also be evidence of a germline variant incorrectly called somatic.

(4) Untested—there were no additional BAMs or there were additional BAMs but none passed the minimum coverage threshold or there were additional BAMs that did not show the variant and so did not provide evidence for or against it.

Telomere length analysis. Reads containing the telomeric repeat (TTAGGG) $\times 3$ or (CCCTAA) $\times 3$ were counted and normalized to the average genomic coverage (the average base coverage of each genome). The normalized telomere count was obtained separately for each tumour and its matching normal. A ratio was calculated by tumour normalized counts/normal normalized counts.

Determination of the BRCA signature. High confidence somatic mutations that were called by both qSNP and GATK across the genome were used to determine the proportion of the BRCA signature in each sample using a published computational framework^{20,49}. In this way, the 96 substitution classification (as determined by substitution class and sequence context) was determined for each sample and compared to the validated BRCA signature²⁰ and the proportion of the BRCA signature in a given sample was ascertained.

Patient derived xenograft (PDX) mouse model generation. Six female eight-week-old NOD/SCID/interleukin 2 receptor [IL2R] gamma (null) (NOG) mice and athymic Balb-c-nude mice were used for the establishment of the patient derived xenograft (PDX) model. All mice were bred at the Australian Bioresources (ABR) under research protocols approved by the Garvan Animal Ethics Committee (09/19, 11/23, 11/09).

The PDXs were generated according to methodology published elsewhere with modifications⁵⁰. Briefly, surgical non-diagnostic specimens of patients operated at APGI clinical sites were implanted subcutaneously (s.c.) into three NOG and three Balb-c-nude mice for each patient, with two small pieces per mouse (left and right flank; engraftment stage). Once established, tumours were grown to a size of 1,500 mm³, at which point they were harvested, divided, and re-transplanted into further mice to bank sufficient tissues for experimentation (first passage and second passage). After expansion, passaged tumours were excised and propagated to cohorts of 40 female Balb-c-nude mice or greater at an average of 8 weeks old, which constituted the treatment cohort (third passage). Utilization of the NOG mouse model, which is characterized by high immune deficiency in this study has enabled establishment of a significant cohort of PDXs (80) xenografts, with a high rate of successful engraftment and propagation (76%, data not shown).

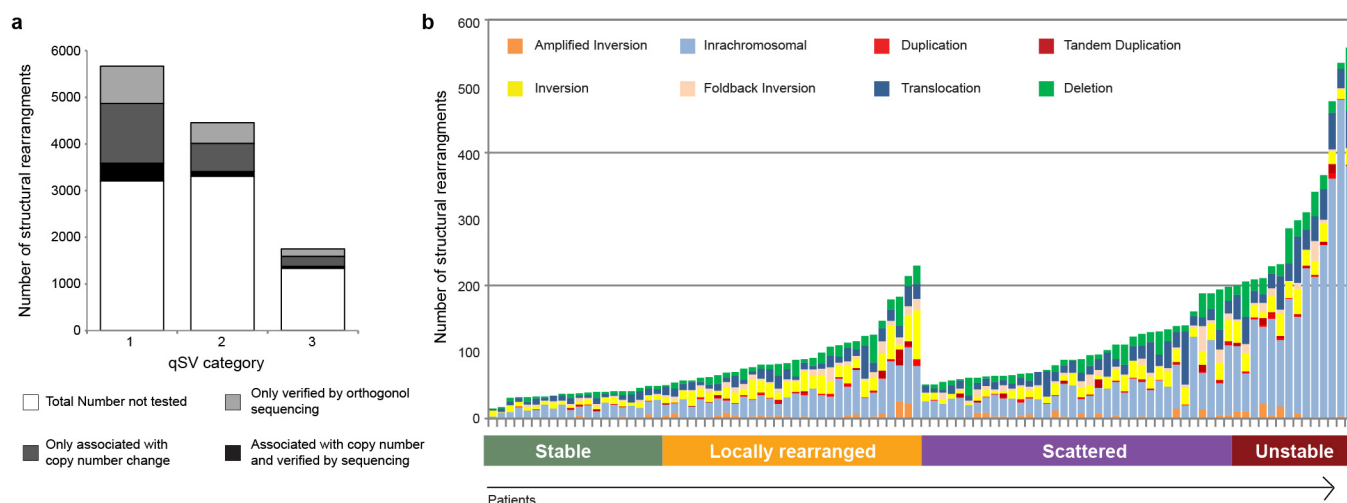
In vivo therapeutic testing. Tumour-bearing mice with a palpable tumour (volume (V) = 150 mm³; V = 0.5 \times length \times width²) were treated with various agents at maximum tolerable dose (MTD) or vehicle treatment based on previously established schedules^{50,51}, where gemcitabine (140 mg per kg) was administered intraperitoneally on day 1 and day 4 for 4 weeks and cisplatin (6 mg per kg) intravenously on day 1 and day 14. The investigators were not blinded to the group allocation. To

avoid accumulating toxicity of repeated injections, an additional treatment was given after the recovery time of two weeks only when no tumour regression was observed, otherwise treatment was continued once the tumour relapsed to its original size (100%). Measurement of chemotherapy response was based on published methodology⁵¹, where primary xenografts were treated with the specified monotherapy and their growth characteristics mapped from the time resistance developed (characterized by progressive tumour growth in the presence of drug), until euthanasia. Mice were euthanized and tissues collected for further analyses when tumour size reached 400% (600–700 mm³).

RAD51 foci formation assay. Antibodies used included RAD51 (Clone 14B4, GeneTex), γ H2AX (phospho-histone H2AX Ser129 clone 20E3, Cell signaling), and geminin (10802-1-AP, ProteinTech Group, Chicago, IL). Primary culture of PDX from patient ICGC_0016 was established by plating and growing cells from an enzymatically digested xenograft on a collagen matrix for approximately 1 week before irradiation and immunofluorescence staining. For this experiment, xenograft was established in a NSG-eGFP mouse. This mouse model allowed us to efficiently visualize eGFP positive mouse stromal cells and eGFP negative tumour cells under the microscope. Briefly, the eGFP expressing NSG mouse was generated in our laboratory by crossing previously established heterozygous eGFPNOD.CB17-Prkdcscid mice⁵² with the theNOD/SCID/interleukin 2 receptor (IL2R) gamma (null) (NOG) strain in our laboratory. eGFP expressing offspring was backcrossed five times onto the parental line to ensure homozygosity for IL2Rgamma deletion and confirmed by genotyping (Transnetyx).

Cell lines of interest were grown on coverslips overnight and irradiated with 10 Gy or left untreated. Subsequently coverslips were fixed with 4% paraformaldehyde (in PBS) 6 h post-irradiation and stained with RAD51, γ H2AX and geminin antibodies as previously described⁵³. DAPI was used as a nuclear stain. RAD51 focus assay scoring was performed as previously established⁵³.

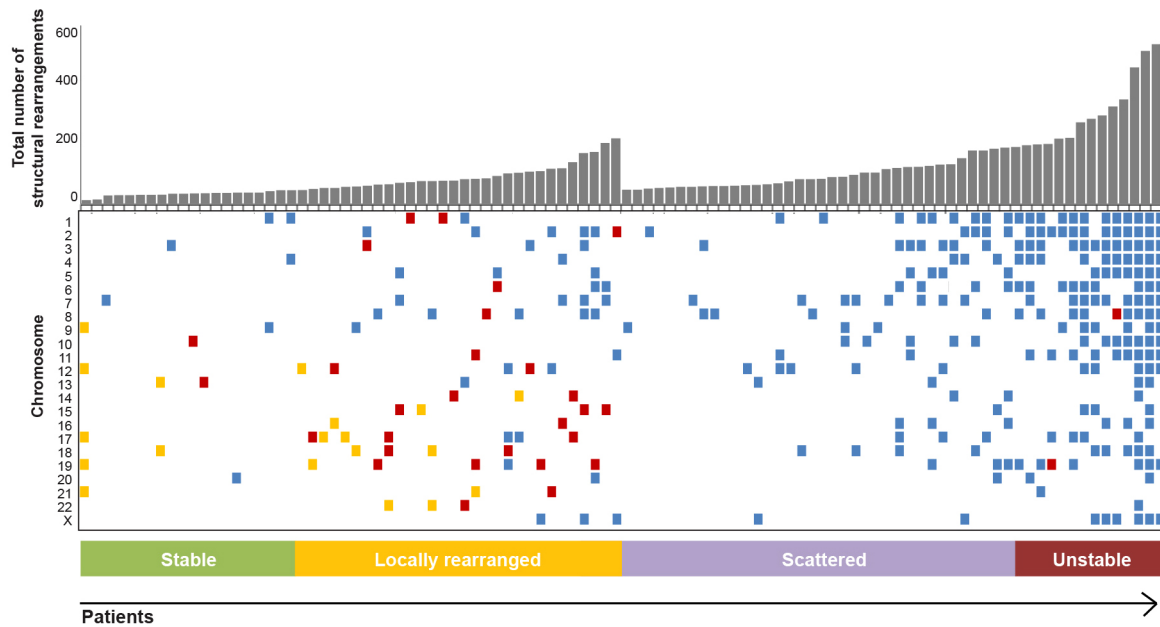
42. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
43. Sun, W. *et al.* Integrated study of copy number states and genotype calls using high-density SNP arrays. *Nucleic Acids Res.* **37**, 5365–5377 (2009).
44. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
45. Nik-Zainal, S. *et al.* Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nature Genet.* **46**, 487–491 (2014).
46. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
47. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).
48. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
49. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
50. Rubio-Viqueira, B. *et al.* An *in vivo* platform for translational drug development in pancreatic cancer. *Clin. Cancer Res.* **12**, 4652–4661 (2006).
51. Rottenberg, S. *et al.* Selective induction of chemotherapy resistance of mammary tumors in a conditional mouse model for hereditary breast cancer. *Proc. Natl Acad. Sci. USA* **104**, 12117–12122 (2007).
52. Niclou, S. P. *et al.* A novel eGFP-expressing immunodeficient mouse model to study tumor-host interactions. *FASEB J.* **22**, 3120–3128 (2008).
53. Graeser, M. *et al.* A marker of homologous recombination predicts pathologic complete response to neoadjuvant chemotherapy in primary breast cancer. *Clin. Cancer Res.* **16**, 6159–6168 (2010).



Extended Data Figure 1 | Summary of structural rearrangements.

a, Histogram showing the number of events verified in silico or by orthogonal sequencing methods (Methods). In total 7,228 of the 11,868 events identified (61%) were verified, the others remain untested. These included 5,666 events which contained multiple lines of evidence (qSV category 1: discordant pairs, soft clipping on both sides and split read evidence, Methods) thus were considered verified. Of these events 2,463 events were also verified by orthogonal sequencing methods (SOLiD long mate pair or PCR amplicon sequencing) or the event was associated with a copy number change which was determined using SNP arrays. The remaining 1,562 events were verified using orthogonal sequencing methods or the event was associated with a copy

number change (qSV category 2 and 3, Methods). **b**, Histogram showing the number of structural rearrangements in each pancreatic cancer. 100 PDACs were sequenced using HiSeq paired-end whole-genome sequencing. Structural rearrangements were identified and classified into 8 categories (deletions, duplications, tandem duplications, foldback inversions, amplified inversions, inversions, intra-chromosomal and inter-chromosomal translocations, Methods). The number and type of event for each patient is shown. PDAC shows a high degree of heterogeneity in both the number and types of events per patient. The structural rearrangements were used to classify the tumours into four categories (stable, locally rearranged, scattered and unstable, Methods).



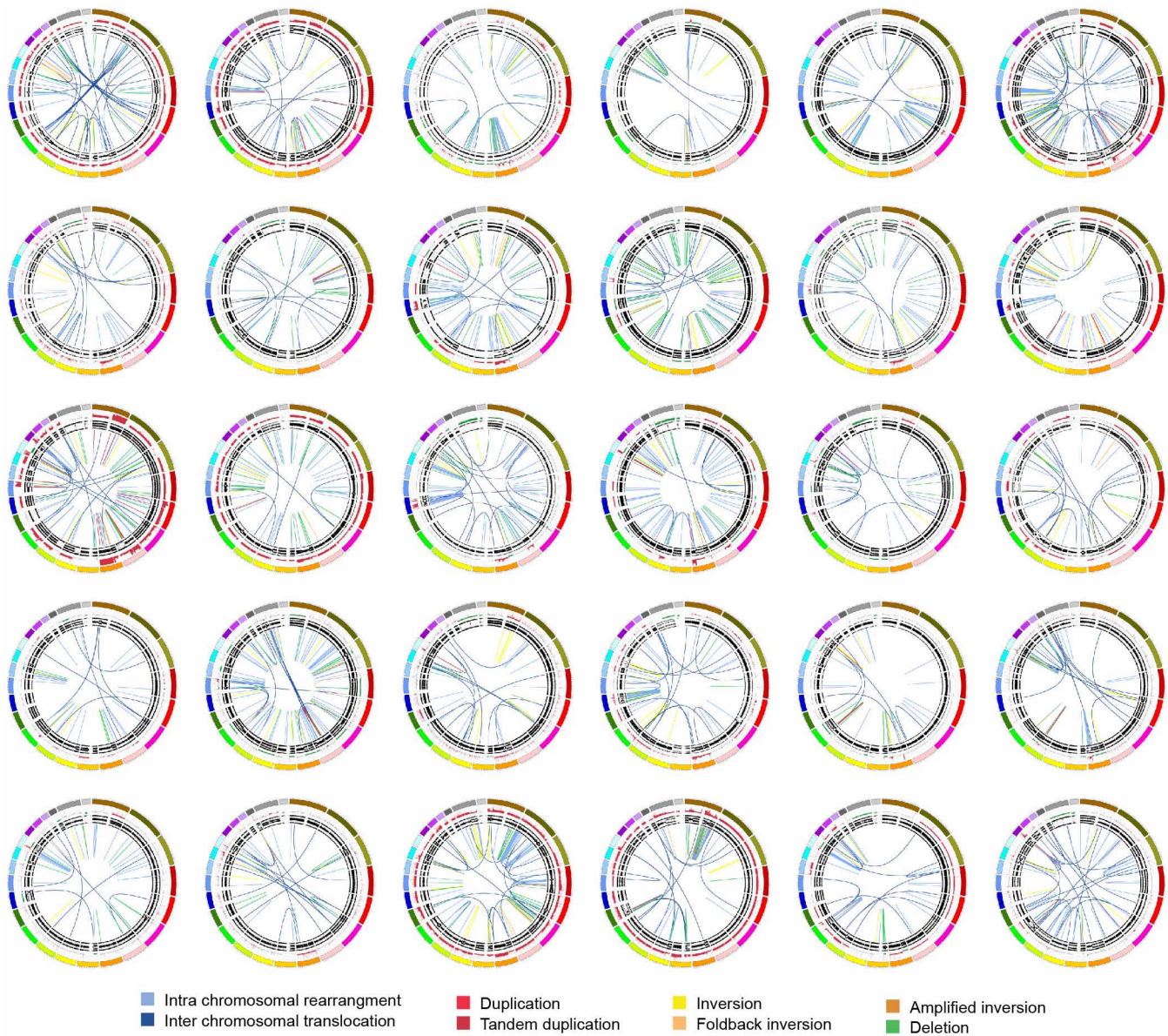
Extended Data Figure 2 | Distribution of structural variant breakpoints within each patient. The 100 patients are plotted along the x axis. The upper plot shows the number of structural rearrangements (y axis) in each tumour. The lower plot shows which chromosomes (y axis) harbour clusters of breakpoints. The distribution of breakpoints (events per Mb) within each chromosome for each sample was evaluated using two methods to identify clusters of rearrangements or chromosomes which contain a large number of events. Method 1: chromosomes with a significant cluster of events were determined by a goodness-of-fit test against the expected exponential distribution (with a significance threshold of <0.0001). Chromosomes which pass these criteria are coloured blue. Method 2: chromosomes were identified

which contain significantly more events per Mb than other chromosomes for that patient. Chromosomes were deemed to harbour a high number of events if they had a mutation rate per Mb which exceeds 1.5 times the length of the interquartile range from the 75th percentile of the chromosome counts for each patient. Chromosomes which pass these criteria are coloured orange. Chromosomes which pass both tests they are coloured red. These criteria show that the unstable tumours which contain many events often have significant clusters of events. In contrast locally rearranged tumours are associated with both clusters of events and a high number of events within that chromosome when compared to other chromosomes.



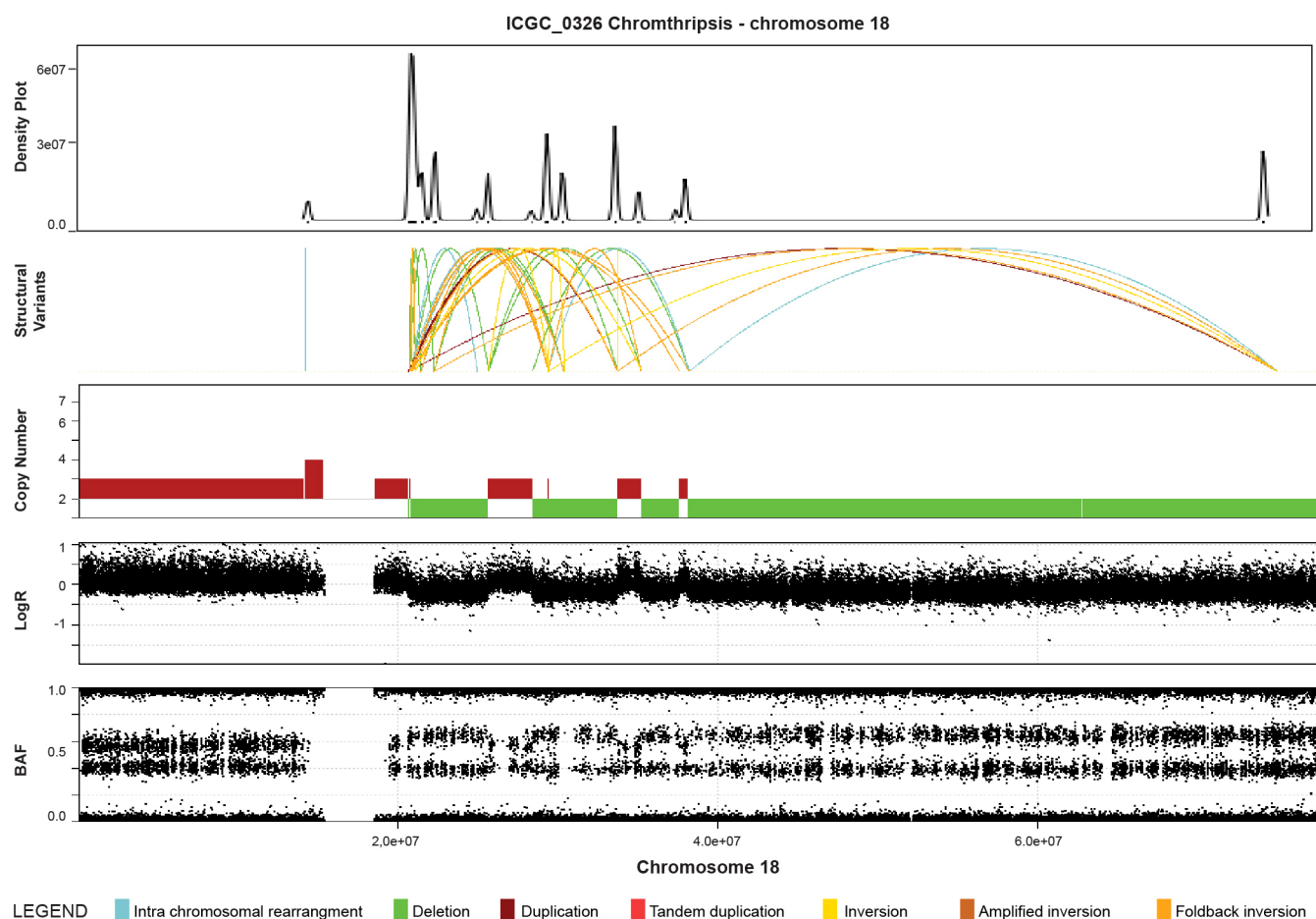
Extended Data Figure 3 | The stable subtype in pancreatic ductal adenocarcinoma. The 20 stable tumours are shown using circos. The coloured outer ring represents the chromosomes, the next ring depicts copy number (red represents gain and green represents loss), the next is the B allele frequency.

The inner lines represent chromosome structural rearrangements detected by whole genome paired sequencing and the legend indicates the type of rearrangement. Stable tumours contained less than 50 structural rearrangements in each tumour.



Extended Data Figure 4 | The locally rearranged subtype in pancreatic ductal adenocarcinoma. The 30 locally rearranged tumours are shown using circos. The coloured outer rings represent the chromosomes, the next ring depicts copy number (red represents gain and green represents loss), the next is

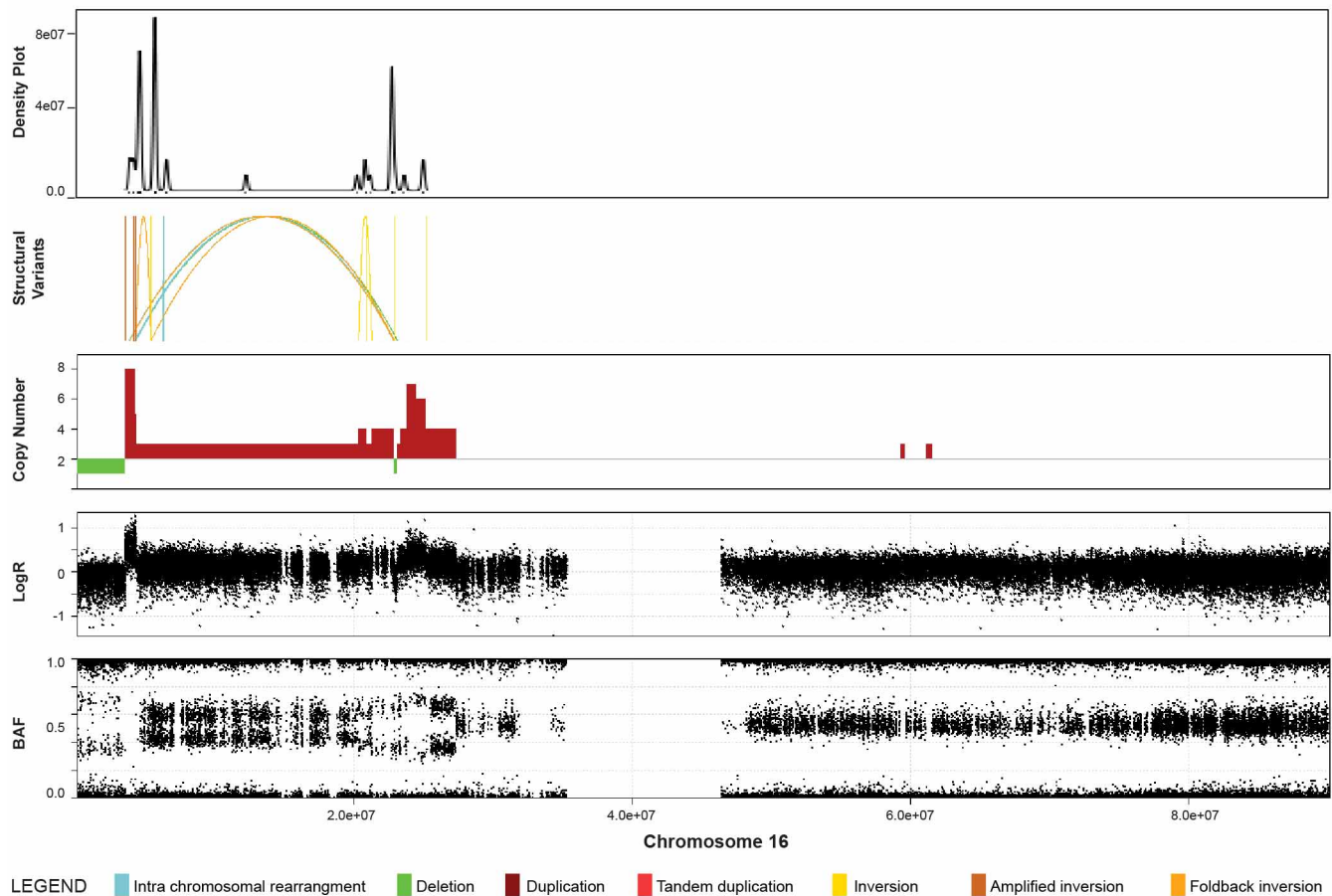
the B allele frequency. The inner lines represent chromosome structural rearrangements detected by whole-genome paired sequencing and the legend indicates the type of rearrangement. In the locally rearranged subtype over 25% of the structural rearrangements are clustered on one of few chromosomes.



Extended Data Figure 5 | Example of evidence for chromothripsis in a pancreatic ductal adenocarcinoma (ICGC_0109). Upper plot is a density plot showing a concentration of break-points on chromosome 5. Next panel shows the structural rearrangements which are coloured as presented in the legend. The lower panels show copy number, logR ratio and B allele frequency derived

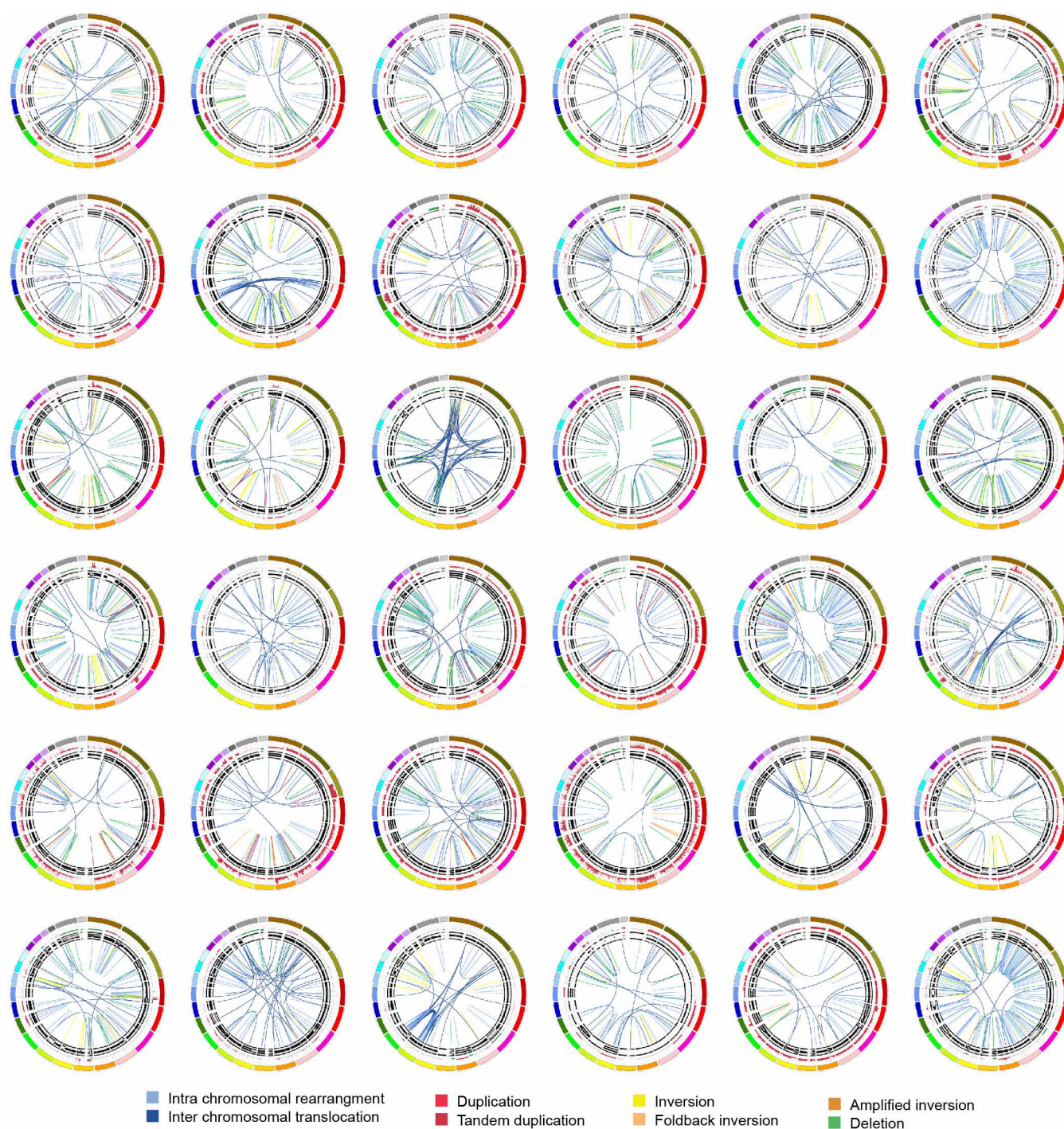
from SNP arrays. This chromosome showed a complex localization of events similar to chromothripsis. Copy number profile and structural rearrangements suggest a shattering of chromosome 5 with a high concentration of structural rearrangements, switches in copy number state and retention of heterozygosity, which are characteristics of a chromothriptic event.

ICGC_0115 Breakage-fusion-bridge - chromosome 16



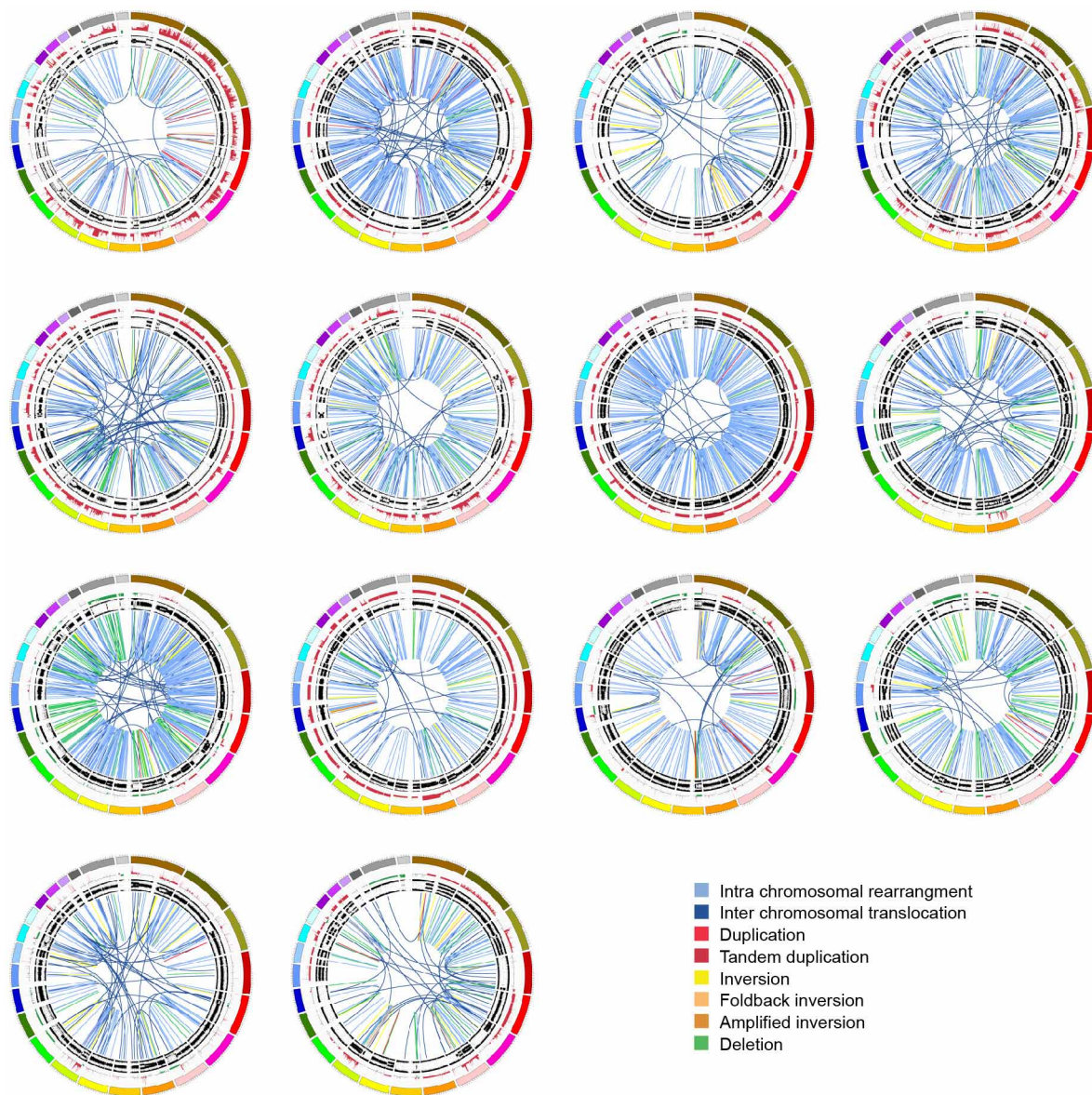
Extended Data Figure 6 | Example of evidence for breakage-fusion-bridge (BFB) in a pancreatic ductal adenocarcinoma (ICGC_0042). Upper plot is a density plot showing a concentration of break-points on chromosome 5. Next panel shows the structural rearrangements which are coloured as presented in the legend. The lower panels show copy number, logR ratio and B

allele frequency derived from SNP arrays. This chromosome showed a complex localization of events similar to BFB. Copy number profile suggests loss of telomeric q arm and a high concentration of structural rearrangements suggesting a series of BFB cycles, with multiple inversions mapped to the amplified regions.



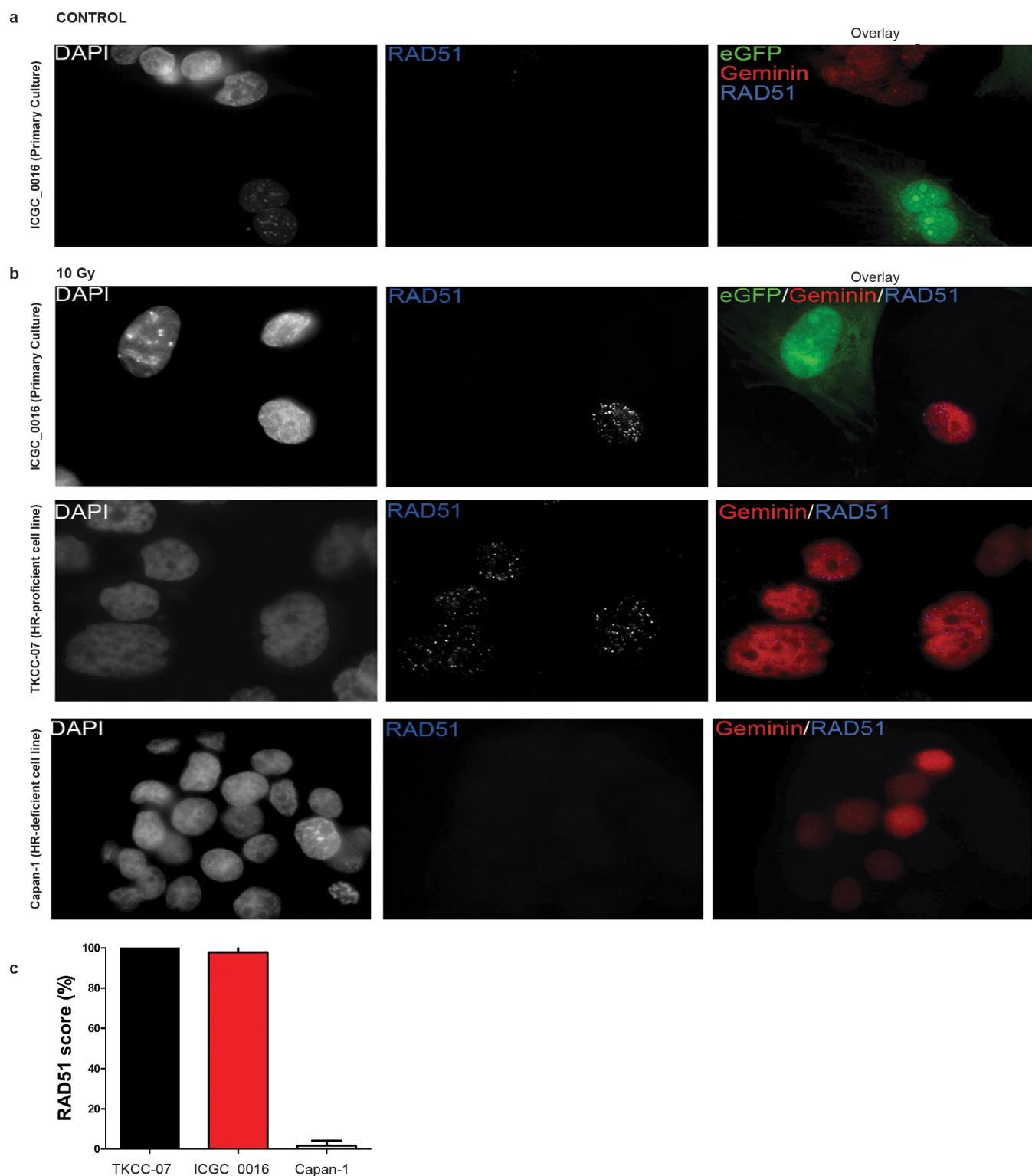
Extended Data Figure 7 | The scattered subtype in pancreatic ductal adenocarcinoma. The 36 tumours classified as scattered are shown using circos. The coloured outer rings represent the chromosomes, the next ring depicts copy number (red represents gain and green represents loss), the next

shows the B allele frequency. The inner lines represent chromosome structural rearrangements detected by whole genome paired end sequencing. The legend indicates the type of rearrangement. The scattered tumours contained 50–200 structural rearrangements in each tumour.



Extended Data Figure 8 | The unstable subtype in pancreatic ductal adenocarcinoma. The 14 unstable tumours are shown using circos. The coloured outer rings are chromosomes, the next ring depicts copy number (red represents gain and green represents loss), the next is the B allele frequency. The inner lines represent chromosome structural rearrangements detected by

whole genome paired sequencing and the legend indicates the type of rearrangement. The unstable tumours contained a large degree of genomic instability and harboured over 200 structural rearrangements in each tumour which were predominantly intra-chromosomal rearrangements evenly distributed through the genome.



Extended Data Figure 9 | RAD51 foci formation in a primary culture of genomically unstable PDAC. **a**, RAD51 and geminin fluorescence in untreated cells derived from an unstable pancreatic tumour with a somatic mutation in the RPA1 gene (ICGC_0016). Primary culture of ICGC_0016 consists of eGFP⁺ mouse stromal and eGFP⁻ tumour cells. **b**, Upper panel: irradiated unstable pancreatic cancer cells (ICGC_0016), middle panel: HR-competent (TKCC-07) and lower panel: HR-deficient (Capan-1) pancreatic tumour cells. Cells were irradiated *in vitro* with 10Gy, and 6 h post-irradiation

examined by immunofluorescence microscopy. eGFP negative tumour cells from ICGC_0016 readily form RAD51 foci following induction of DNA damage. TKCC-07 is a pancreas cancer cell line generated from a homologous recombination (HR) pathway competent patient-derived xenograft and served as a positive control for staining and RAD51 foci formation after DNA damage. Capan-1 cells which are HR-deficient do not form RAD51 foci. **c**, RAD51 score (percentage of geminin positive cells that have RAD51 foci) in examined pancreatic tumour cells.

Lagging-strand replication shapes the mutational landscape of the genome

Martin A. M. Reijns^{1*}, Harriet Kemp^{2*}, James Ding¹, Sophie Marion de Procé², Andrew P. Jackson¹ & Martin S. Taylor²

The origin of mutations is central to understanding evolution and of key relevance to health. Variation occurs non-randomly across the genome, and mechanisms for this remain to be defined. Here we report that the 5' ends of Okazaki fragments have significantly increased levels of nucleotide substitution, indicating a replicative origin for such mutations. Using a novel method, emRiboSeq, we map the genome-wide contribution of polymerases, and show that despite Okazaki fragment processing, DNA synthesized by error-prone polymerase- α (Pol- α) is retained *in vivo*, comprising approximately 1.5% of the mature genome. We propose that DNA-binding proteins that rapidly re-associate post-replication act as partial barriers to Pol- δ -mediated displacement of Pol- α -synthesized DNA, resulting in incorporation of such Pol- α tracts and increased mutation rates at specific sites. We observe a mutational cost to chromatin and regulatory protein binding, resulting in mutation hotspots at regulatory elements, with signatures of this process detectable in both yeast and humans.

Mutations occur despite the exquisite fidelity of DNA replication, efficient proofreading and mismatch repair¹, resulting in heritable disease and providing the raw material for evolution. Genome variation is non-uniform², the outcome of diverse mutational processes³, repair mechanisms⁴ and selection pressures^{5,6}. This variability is exemplified by nucleotide substitution rates around nucleosome binding sites, with the highest rates at the nucleosome midpoint (dyad position)^{7–12}.

Bidirectional replication of genomic DNA necessitates discontinuous synthesis of the lagging strand as a series of Okazaki fragments (OFs)^{13,14}, which then undergo processing to form an intact continuous DNA strand^{15,16}. Recently, the genomic locations at which OFs are ligated (Okazaki junctions, OJs) were mapped¹⁷. In this experimental system, OJs occurred at an average rate of 0.6% per nucleotide; however, frequency was strongly influenced by the binding of nucleosomes and transcription factors (TFs). These proteins act as partial blocks to Pol- δ processivity, resulting in the accumulation of OJs at their binding sites. Here, we demonstrate the mutational consequences of such protein binding.

Substitutions correlate with OJs

We were struck by the similarity of the distribution of *Saccharomyces cerevisiae* OJ sites at nucleosomes¹⁷ to that previously reported for nucleotide substitutions^{7,8,10–12}, and set out to investigate the potential reasons for this. We established that nucleotide substitution and OJ distributions are highly correlated (Pearson's correlation coefficient = 0.76, $P = 2.2 \times 10^{-16}$) and essentially identical in pattern (Fig. 1a). Furthermore, differences in OJ distribution by nucleosome type (genic versus non-genic), spacing or consistency of binding were mirrored by the substitution rate distribution (Extended Data Fig. 1a–f). We found similar strong correlation in the regions directly surrounding TF binding sites of Reb1 (Fig. 1b; Pearson's correlation = 0.57, $P = 5.6 \times 10^{-15}$) and Rap1 (Extended Data Fig. 1g), providing further evidence for a direct association. At the sequence-specific binding sites themselves, substitution rates were depressed relative to the OJ, resulting from strong

selection pressure to maintain TF binding, and obscuring any mutational signal at these nucleotides.

Given that both classes of sites (nucleosomes and TFs) are present genome-wide and represent different biological processes, this association was probably the direct consequence of protein binding at these sites. However, to rule out site-specific biases in sequence as a confounding explanation for the observed distributions, we randomly sampled the rest of the genome for trinucleotides of identical sequence compositions and calculated the substitution rate at these sites, on a nucleotide-by-nucleotide position basis (Extended Data Fig. 1h–j). This resulted in loss of the observed patterns, establishing that nucleotide composition bias was not a contributing factor. Furthermore, the observed association was not restricted to polymorphism rates, as yeast inter-species nucleotide substitution patterns at both nucleosome and Reb1 TF binding sites were identical (Extended Data Fig. 1k, l).

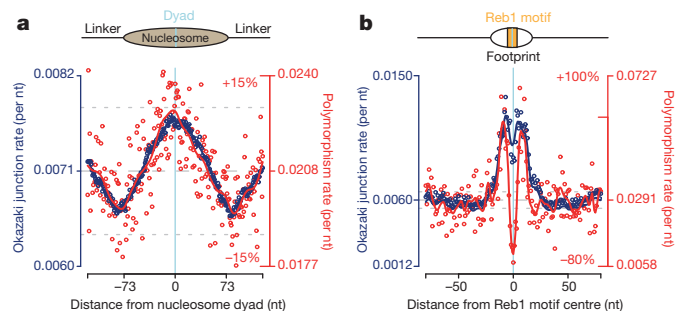


Figure 1 | Increased substitution rates at OJs. **a, b**, Nucleotide (nt) substitution rates (red) closely correlate with increased OJ site frequency (blue) at nucleosome (**a**) and Reb1 (**b**) binding sites. *S. cerevisiae* polymorphism rates per nucleotide computed using sequences from nucleosome ($n = 27,586$) and Reb1 binding sites ($n = 881$). Individual data points, open circles. Solid curves, best-fit splines. Mean, dashed grey line; $\pm 10\%$ dotted grey lines.

¹Medical and Developmental Genetics, MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. ²Biomedical Systems Analysis, MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK.

*These authors contributed equally to this work.

We therefore concluded that OJ frequency and nucleotide substitution rates could be causally related, and set out to investigate the potential mechanism for this association.

Mutations at 5' ends of OFs

The synthesis and processing of OFs is directional. Therefore, substitution rates would be expected to be asymmetrical relative to the direction of synthesis, if a component of this process was the cause. As most of the genome is preferentially replicated with either the forward or reverse strand as the lagging strand, we orientated regions by their dominant direction of lagging-strand synthesis. This revealed substantially increased nucleotide substitution rates immediately downstream of OJs (Fig. 2a), the level of mutational signal correlating with OJ site frequency. Quantification of substitution rates for the five nucleotides immediately upstream and downstream of the OJ (Fig. 2b) demonstrated that high frequency OJ sites (11-fold increased OJ rate relative to baseline; top 99.9th centile of sites) displayed the highest substitution rate ($P < 2.2 \times 10^{-16}$), with significant increases ($P < 2.2 \times 10^{-16}$) for medium frequency sites, (6.1-fold, 99–99.9th centile) but not low frequency sites ($P = 0.3$, 1.7-fold, OJ sites <99th centile). This was not due to site-specific sequence biases, as the increase in substitution rate was lost after a trinucleotide preserving genome shuffle. Therefore, point mutations are enriched at the 5' ends of mature OFs of frequently occurring OJ sites, sites that correspond to protein barriers to Pol- δ processivity¹⁷.

Pol- α DNA retention hypothesis

We next considered which aspect of lagging-strand synthesis might be responsible. OFs are generated by the consecutive actions of Pol- α and Pol- δ (Fig. 2c). When the previously synthesized, downstream OF is encountered, OF processing occurs¹⁸, involving the coordinated action of FEN1 and DNA2 nucleases^{15,16} in conjunction with continuing DNA synthesis by Pol- δ , before final ligation of adjoining DNA fragments. During this process, most if not all of the 10–30-nucleotide-long DNA primer synthesized by Pol- α ^{19,20} has been thought to be removed alongside the RNA primer, and replaced by Pol- δ -synthesized DNA^{16,21,22}.

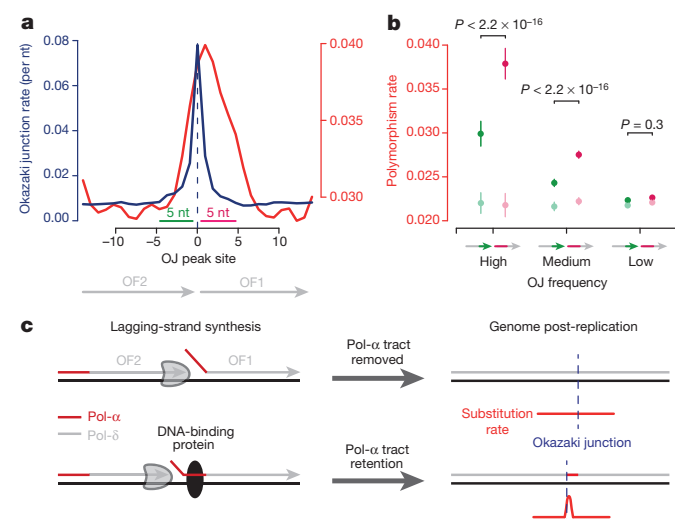


Figure 2 | Frequent nucleotide substitutions at OF 5' ends. **a**, Mutation rates are increased downstream of OJs. Substitution polymorphisms (red) and OJ rate (blue) in regions surrounding high frequency OJs (top 0.1%). $n = 5,660$ sequences orientated for dominant direction of OF synthesis. **b**, Mutation rates correlate with OJ peak size. Mutations are significantly enriched downstream of the junction (pink), compared to genome shuffle controls (light green/pink). Sites grouped by OJ frequency. Points denote mean and error bars denote s.d. from 100 bootstrap samples or genome shuffles (controls); statistics by paired two-sided t -test. **c**, Hypothesis: DNA synthesized by non-proofreading Pol- α is preferentially trapped in regions rapidly bound by proteins post-replication. These act as partial barriers to Pol- δ displacement of Pol- α -synthesized DNA, resulting in locally increased mutations.

This would be desirable, as unlike other replicative DNA polymerases, Pol- α lacks 3'-to-5' proofreading exonuclease activity, limiting its intrinsic fidelity²³. On the other hand, studies on the mutagenesis pattern of reduced fidelity polymerase mutants in yeast demonstrate that Pol- α -synthesized DNA does contribute to the genome^{21,24–26}. How comprehensive the removal or retention of such DNA is *in vivo* is unknown, but notably the retention of error-prone Pol- α -synthesized DNA at the 5' end of OFs would provide a straightforward explanation for the increased mutation rates we observed. Given that protein barriers have been shown to influence OF processing¹⁷, we therefore propose that Pol- α -synthesized DNA is preferentially retained at sites where proteins bind shortly after initial OF DNA synthesis (Fig. 2c). Our model would predict (1) that Pol- α tracts are retained at a considerable level within the mature genome post-replication, and (2) that mutational signatures arising from such Pol- α -synthesized DNA will be increased at many DNA-binding protein sites in eukaryotes.

EmRiboSeq

To address where error-prone Pol- α DNA is retained *in vivo*, we used the incorporation of ribonucleotides into genomic DNA to track the activity of specific DNA polymerases. Ribonucleotides are covalently incorporated into genomic DNA by replicative polymerases^{27,28}, although they are normally efficiently removed by ribonucleotide excision repair, a process initiated by the type 2 RNase H enzyme (RNase H2)²⁹. In RNase-H2-deficient budding yeast, such ribonucleotides are generally well tolerated: $\Delta rnh201$ yeast has proliferation rates identical to wild type under normal growth conditions²⁷, and therefore in this genetic background ribonucleotides can be used as a 'label' to track polymerase activity. Furthermore, the contribution of specific polymerases can be studied using polymerases with catalytic site point mutations (such as Pol- α (Leu868Met), Pol- δ (Leu612Met) and Pol- ϵ (Met644Gly)) that incorporate ribonucleotides at higher rates than their wild-type counterparts (refs 21, 26, 27, 30 and J. S. Williams, A. R. Clausen & T. A.

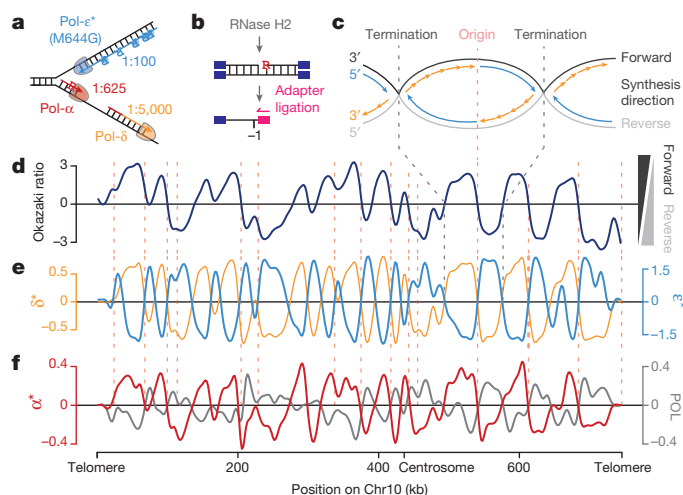


Figure 3 | Mapping DNA synthesis *in vivo* using emRiboSeq. **a**, Replicative polymerases can be tracked using point mutants with increased ribonucleotide incorporation. Schematic of replication fork with Pol- ϵ (asterisk denotes Met644Gly mutant) and ribonucleotide incorporation rates for each polymerase. Embedded ribonucleotides (R) highlighted. **b**, Schematic of emRiboSeq methodology. **c**, Schematic of replication. **d**, **e**, Mapping of leading/lagging-strand synthesis and replication origins using emRiboSeq. Ratio of OF reads¹⁷ between forward and reverse strands of chromosome 10 (Chr10; **d**) corresponds to the ratio of their respective ribonucleotide content (**e**) for Pol- δ * (orange), whereas Pol- ϵ * (cyan) shows negative correlation. Intersections with x axis correspond to replication origins and termination regions (**c–e**). Experimentally validated origins (dotted pink lines). **f**, Pol- α * DNA is detected genome-wide by emRiboSeq as a component of the lagging strand. Strand ratios are shown as best-fit splines, y axes denote \log_2 of ratios (**d–f**).

Kunkel, personal communication; Fig. 3a). Yeast strains expressing these mutant polymerases have previously been used to demonstrate that Pol- ϵ and Pol- δ are the major leading- and lagging-strand polymerases, respectively, by measuring strand-specific alkaline sensitivity of particular genomic loci^{30–32}.

To track directly the genome-wide contribution of polymerases, we developed a next-generation sequencing approach, which we term emRiboSeq (for embedded ribonucleotide sequencing), that determines the strand-specific, genome-wide distribution of embedded ribonucleotides. This is achieved by treatment of genomic DNA with recombinant RNase H2 to generate nicks 5' of embedded ribonucleotides, followed by ligation of a sequencing adaptor to the 3'-hydroxyl group of the deoxynucleotide immediately upstream of the ribonucleotide (Fig. 3b and Extended Data Fig. 2a). Subsequent ion-semiconductor sequencing permits strand-specific mapping of ribonucleotide incorporation sites.

Control experiments using endonucleases of known sequence specificity demonstrated 99.9% strand specificity and 99.9% site specificity for the technique (Extended Data Fig. 2b–d). Using RNase-H2-deficient Pol- ϵ (Met644Gly) and Pol- δ (Leu612Met) yeast strains, we then mapped the relative contributions of these respective polymerases genome-wide (Fig. 3c–e and Extended Data Figs 3 and 4). We found that ribonucleotide incorporation in the Pol- δ (Leu612Met) strain was substantially enriched on the DNA strand that is preferentially synthesized by lagging-strand synthesis¹⁷, in keeping with its function as the major lagging-strand polymerase^{30,33,34}, while ribonucleotide incorporation in the Pol- ϵ (Met644Gly) strain exhibited an entirely reciprocal pattern consistent with its function as the leading-strand polymerase^{31,35} (Fig. 3e). Furthermore, points at which neither enzyme showed strand preference (intersection of both Pol- ϵ and Pol- δ plots with the x axis) corresponded precisely with annotated origins of replication. Other intersection points were also evident that correspond to replication termination regions, as well as putative, non-annotated origins. The latter overlapped with early replicating regions³⁶ (Extended Data Fig. 3b, c). Therefore, we concluded that emRiboSeq can be used to determine the distribution of polymerase activity genome-wide, and has utility for the identification of replication origin and termination sites.

Pol- α -synthesized DNA ~1.5% of genome

Having demonstrated the validity of our technique through detailed mapping of the major replicative polymerases, we next examined the contribution of Pol- α -synthesized DNA to the budding yeast genome. Significantly, the Pol- α (Leu868Met) Δ rnh201 strain had a strand ratio distribution identical to that seen for Pol- δ (Leu612Met) Δ rnh201, consistent with the expected role for Pol- α in lagging-strand replication (Fig. 3f). Furthermore, the Pol- α (Leu868Met) pattern of strand incorporation was reciprocal to that of a wild-type polymerase strain (POL), which displayed leading-strand bias, in keeping with a strong propensity for ribonucleotide incorporation by leading-strand polymerase Pol- ϵ compared to Pol- δ (ref. 37). Increased ribonucleotide retention on the lagging strand was also present in DNA from stationary phase Pol- α (Leu868Met) Δ rnh201 yeast (Extended Data Fig. 3d), demonstrating that Pol- α -derived DNA is retained in the mature genome post-replication and that this signal was not due to the transient presence of Pol- α DNA during S-phase.

To provide biochemical validation, we performed alkaline gel electrophoresis on genomic DNA extracted from Pol- α (Leu868Met), Pol- δ (Leu612Met) and Pol- ϵ (Met644Gly) Δ rnh201 yeast. Increased fragmentation was detected in all three strains (Extended Data Fig. 4a–c) and increased ribonucleotide incorporation was also detected in genomic DNA from stationary phase Pol- α (Leu868Met) yeast (Fig. 4a–c), consistent with Pol- α tract retention in mature genomic DNA. To quantify the contribution of Pol- α DNA to the genome, we used densitometry measurements from the alkaline gels to calculate ribonucleotide incorporation rates²⁸. We detected 1,500 embedded ribonucleotides per genome in Δ rnh201 genomic DNA, which increased to 2,400 sites per genome for Pol- α (Leu868Met) (Fig. 4c). Observed ribonucleotide

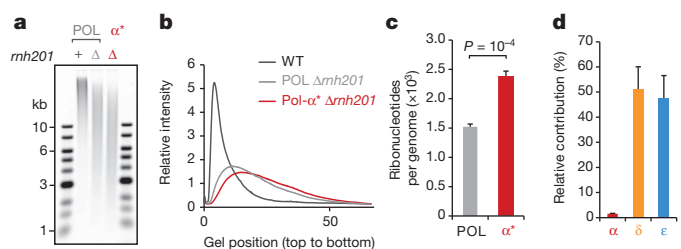


Figure 4 | Pol- α DNA synthesis contributes ~1.5% of the mature genome.

a, b, Increased ribonucleotide incorporation in Pol- α^* stationary phase yeast is detected by alkaline gel electrophoresis. kb, kilobases; WT, wild type.

c, Quantification confirms significantly increased rates in the Pol- α^* genome ($n = 6$ independent experiments; error bars denote s.e.m.; statistics by paired two-sided t -test). **d,** Estimate of relative contribution of polymerases to the genome ($n = 4$ independent experiments; error bars denote s.e.m.).

incorporation rates correspond to the product of the incorporation frequency of each polymerase and the amount of DNA it contributes to the genome. Using the *in vitro* ribonucleotide incorporation rates of wild-type and mutant polymerases and the number of embedded ribonucleotides embedded *in vivo* (Extended Data Figs 3a and 4a–c), we estimated the relative contributions of each of the replicative polymerases to the genome (Fig. 4d), calculating the contribution of Pol- α to be $1.5 \pm 0.3\%$ (mean \pm s.d.).

RNase H enzymes may contribute to the removal of OF RNA primers^{16,38} and consequently Δ rnh201 strains could have altered levels of Pol- α -synthesized DNA to that seen in wild-type strains. This confounding factor was excluded using an RNH201 separation-of-function mutant³⁹, which established that retention of Pol- α DNA was independent of a role for RNase H2 in RNA primer removal (Extended Data Fig. 5).

In conclusion, Pol- α -synthesized DNA makes a small but significant contribution to the genome, relative to the major replicative polymerases, confirming the first prediction of our model.

Mutational cost of TF binding in humans

As OF processing is a conserved process in eukaryotes, we next considered whether an OF-related mutational signature was also present in humans. Substitution rates are also increased at nucleosome cores in humans⁷ with an identical distribution to yeast. Furthermore, the TF NFYA has an unexplained ‘shoulder’ of increased substitution proximal to its binding sites⁴⁰, reminiscent of the Reb1 pattern (Fig. 1b). We therefore investigated whether similar mutational patterns are present at other experimentally defined human TF and chromatin protein binding sites. Increased inter-species nucleotide substitution rates were detected flanking essential binding site residues, for many, but not all TFs, as well as CTCF binding sites (Fig. 5a, b and Extended Data Fig. 6). Substitution rates were measured using genomic evolutionary rate profiling (GERP) scores, which quantify nucleotide substitution rates relative to a genome-wide expectation of neutral evolution⁴¹, such that a negative GERP score indicates increased nucleotide substitution rates. Furthermore, increases in mutation rate correlated with the degree of enrichment reported in chromatin immunoprecipitation with lambda exonuclease digestion (ChIP-exo) data sets for these proteins, likely reflecting the strength of binding or frequency of occupancy at specific sites, which would be expected to influence Pol- δ processivity and consequent mutation levels.

Finally, to extend our analysis beyond common TF binding sites, we investigated whether the same mutational signature could be found for a broad range of regions at which regulatory proteins bind, regions we identified by the presence of DNase I footprints. Our preceding analysis of TFs suggested that nucleotide substitutions would be increased immediately adjacent to the protein binding region defined by such footprints. In yeast we found that DNase I footprint edges served as a good proxy for increased OJ rate with significantly elevated substitution rates

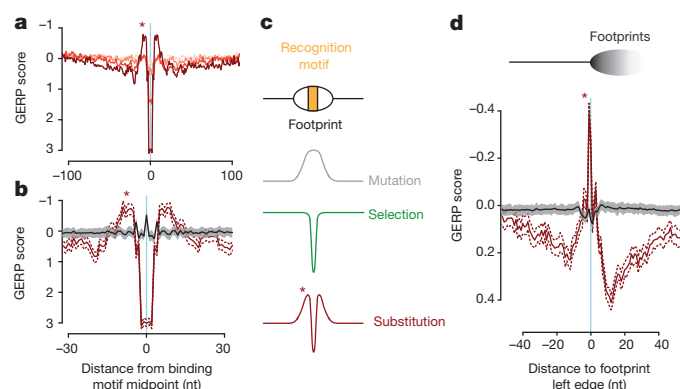


Figure 5 | OF mutational signatures are conserved in humans. **a**, Nucleotide substitutions (plotted as GERP scores) are increased immediately adjacent to TF NFYA binding sites ($n = 5,110$). Pink to brown: lower to higher quartiles of ChIP-exo peak height (reflecting strength of binding/occupancy). Stronger binding correlates with substitution rate in the 'shoulder' region (asterisk). **b**, Increased substitution rates are not a consequence of local sequence composition effects. Strongest binding sites (brown) compared to trinucleotide preserving shuffle (black). **c**, Model showing nucleotide substitution profiles are the sum of mutation rate and selective pressure. **d**, Interspecies substitution rates are also increased adjacent to DNase I footprint edges (asterisk) ($n = 33,350$). Sequences aligned to left footprint edges as indicated in schematic. Right footprint edge is indistinct owing to heterogeneity in footprint length. Substitution rates are no longer increased after trinucleotide preserving shuffle from local flanking sequences (black). Brown dashes and grey shading denote 95% confidence intervals (**b**, **d**).

(Extended Data Fig. 7). Similarly, in humans, aligning regions containing DNase I footprints on the basis of boundary junctions (left-hand edge of footprint), detected substantially increased nucleotide substitution rates close to the junction, relative to the baseline rate in the immediate region (Fig. 5d). These increased substitution rates were related to position rather than sequence content, as this signal was lost when a trinucleotide preserving genome shuffle was applied, both for individual TFs (Fig. 5b and Extended Data Fig. 6a–d) and DNase I footprints (Fig. 5d). Therefore, this mutational signature is not due to the retention of mutagenic sequences (for example, CpG dinucleotides) at such sites⁴², and is a widespread phenomenon in the genome at protein binding sites in both yeast and humans.

Discussion

Here we establish a mutational signature at protein binding sites that we suggest could result from the activity of the replicative polymerase Pol- α . We use a novel technique, emRiboSeq, to demonstrate that error-prone DNA synthesized by Pol- α is retained in the mature lagging strand. EmRiboSeq tracks genome-wide *in vivo* polymerase activity using ribonucleotides as a 'non-invasive' label, and will have significant future use for the *in vivo* study of DNA polymerases in replication and repair. Further optimization of emRiboSeq should permit high resolution examination of the role of polymerases at specific sites, such as Pol- α tract retention at protein binding sites. It will also be a useful method for defining replication origin and termination sites, and furthermore will facilitate the investigation of physiological roles of genome-embedded ribonucleotides^{30,43,44}.

A direct relationship between OF junctions and mutation frequency is indicated by the significant correlations between substitution rate and OF junction sites at diverse protein binding sites, although future experimental validation will be needed to establish causality formally. We find that substitution rates are specifically increased downstream of such junction sites, suggesting a replicative origin for such mutations. As Pol- α DNA tracts occur genome-wide, and Pol- δ processing of OFs is impaired by DNA-bound proteins¹⁷, we propose that retention of Pol- α DNA is increased at these functionally important sites, and is responsible for the increased mutation rate (Fig. 2c). Replication fidelity

processes, including efficient mismatch repair at the 5' end of OFs^{25,45}, will mitigate Pol- α replication errors. Additionally, Pol- α DNA will be incorporated at relatively low frequency (Extended Data Fig. 8), with most DNA at such sites still synthesized by Pol- δ and Pol- ϵ . However, over evolutionary timescales, it seems that these processes are insufficient to compensate fully for the lack of Pol- α proofreading activity. An alternative possibility is that protein binding may impair access of replication-related repair factors, such as Exo1 to correct errors in Pol- α -synthesized DNA⁴⁵. However, it does not appear that the mismatch repair machinery is generally obstructed at such sites, as mismatch repair efficiency at nucleosomes is reported to be uniform with respect to dyad position²⁴.

Nucleosome formation has a key role in ensuring genome stability⁴⁶, and consequently there is an imperative for the rapid repackaging of the genome post-replication. However, we now show that this comes at the cost of increased mutation at specific sites, detectable on an evolutionary timescale. OF-associated mutagenesis could also have importance for human genetics, as it increases mutation rates at TF and regulatory protein binding sites. Such increased mutagenesis has been substantially obscured by strong purifying selection at these sites necessary to maintain functionality. Notably, increased mutation suggests that they will be evolutionary hotspots, and may help to explain the rapid evolutionary turnover of TF sites⁴⁷ and the difficulty in non-coding functional site prediction by interspecies sequence conservation comparisons. Furthermore, as hyper-mutable loci, TF binding sites may be frequently mutated in inherited disease and neoplasia.

In summary, we demonstrate that DNA synthesized by Pol- α contributes to the eukaryotic genome, probably increasing mutations at specific regulatory sites of relevance to both human genetics and the shaping of the genome during evolution.

Note added in proof: Three studies, published concurrently with this paper, have independently developed similar methods to determine the genome-wide distribution of embedded ribonucleotides^{48,49,50}, demonstrating the utility of ribonucleotides as markers of replication enzymology in budding yeast.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 October 2014; accepted 5 January 2015.

Published online 26 January 2015.

- Kunkel, T. A. Evolving views of DNA replication (in)fidelity. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 91–101 (2009).
- Wolfe, K. H., Sharp, P. M. & Li, W. H. Mutation rates differ among regions of the mammalian genome. *Nature* **337**, 283–285 (1989).
- Alexandrov, L. B. & Stratton, M. R. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. *Curr. Opin. Genet. Dev.* **24**, 52–60 (2014).
- Ciccia, A. & Elledge, S. J. The DNA damage response: making it safe to play with knives. *Mol. Cell* **40**, 179–204 (2010).
- Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482 (2011).
- Pollard, K. S. *et al.* Forces shaping the fastest evolving regions in the human genome. *PLoS Genet.* **2**, e168 (2006).
- Prendergast, J. G. & Semple, C. A. Widespread signatures of recent selection linked to nucleosome positioning in the human lineage. *Genome Res.* **21**, 1777–1787 (2011).
- Sasaki, S. *et al.* Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science* **323**, 401–404 (2009).
- Semple, C. A. & Taylor, M. S. Molecular biology. The structure of change. *Science* **323**, 347–348 (2009).
- Warnecke, T., Batada, N. N. & Hurst, L. D. The impact of the nucleosome code on protein-coding sequence evolution in yeast. *PLoS Genet.* **4**, e1000250 (2008).
- Washietl, S., Machne, R. & Goldman, N. Evolutionary footprints of nucleosome positions in yeast. *Trends Genet.* **24**, 583–587 (2008).
- Ying, H., Epps, J., Williams, R. & Huttley, G. Evidence that localized variation in primate sequence divergence arises from an influence of nucleosome placement on DNA repair. *Mol. Biol. Evol.* **27**, 637–649 (2010).
- Johnston, L. H. & Nasmyth, K. A. *Saccharomyces cerevisiae* cell cycle mutant *cdc9* is defective in DNA ligase. *Nature* **274**, 891–893 (1978).
- Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K. & Sugino, A. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc. Natl Acad. Sci. USA* **59**, 598–605 (1968).

15. Balakrishnan, L. & Bambara, R. A. Okazaki fragment metabolism. *Cold Spring Harb. Perspect. Biol.* **5**, a010173 (2013).
16. Zheng, L. & Shen, B. Okazaki fragment maturation: nucleases take centre stage. *J. Mol. Cell Biol.* **3**, 23–30 (2011).
17. Smith, D. J. & Whitehouse, I. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* **483**, 434–438 (2012).
18. Stith, C. M., Sterling, J., Resnick, M. A., Gordenin, D. A. & Burgers, P. M. Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. *J. Biol. Chem.* **283**, 34129–34140 (2008).
19. Perera, R. L. *et al.* Mechanism for priming DNA synthesis by yeast DNA Polymerase α . *Elife* **2**, e00482 (2013).
20. Walsh, E. & Eckert, K. A. Eukaryotic Replicative DNA Polymerases. *Nucleic Acid Polymerases* **30**, 17–41 (2014).
21. Pavlov, Y. I. *et al.* Evidence that errors made by DNA polymerase α are corrected by DNA polymerase δ . *Curr. Biol.* **16**, 202–207 (2006).
22. Maga, G. *et al.* Okazaki fragment processing: modulation of the strand displacement activity of DNA polymerase δ by the concerted action of replication protein A, proliferating cell nuclear antigen, and flap endonuclease-1. *Proc. Natl Acad. Sci. USA* **98**, 14298–14303 (2001).
23. Kunkel, T. A., Hamatake, R. K., Motto-Fox, J., Fitzgerald, M. P. & Sugino, A. Fidelity of DNA polymerase I and the DNA polymerase I-DNA primase complex from *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **9**, 4447–4458 (1989).
24. Lujan, S. A. *et al.* Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res.* **24**, 1751–1764 (2014).
25. Nick McElhinny, S. A., Kissling, G. E. & Kunkel, T. A. Differential correction of lagging-strand replication errors made by DNA polymerases α and δ . *Proc. Natl Acad. Sci. USA* **107**, 21070–21075 (2010).
26. Niimi, A. *et al.* Palm mutants in DNA polymerases α and η alter DNA replication fidelity and translesion activity. *Mol. Cell Biol.* **24**, 2734–2746 (2004).
27. Nick McElhinny, S. A. *et al.* Genome instability due to ribonucleotide incorporation into DNA. *Nature Chem. Biol.* **6**, 774–781 (2010).
28. Reijns, M. A. *et al.* Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* **149**, 1008–1022 (2012).
29. Sparks, J. L. *et al.* RNase H2-initiated ribonucleotide excision repair. *Mol. Cell* **47**, 980–986 (2012).
30. Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B. & Kunkel, T. A. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol. Cell* **50**, 437–443 (2013).
31. Lujan, S. A. *et al.* Mismatch repair balances leading and lagging strand DNA replication fidelity. *PLoS Genet.* **8**, e1003016 (2012).
32. Miyabe, I., Kunkel, T. A. & Carr, A. M. The major roles of DNA polymerases ϵ and δ at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet.* **7**, e1002407 (2011).
33. Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M. & Kunkel, T. A. Division of labor at the eukaryotic replication fork. *Mol. Cell* **30**, 137–144 (2008).
34. Larrea, A. A. *et al.* Genome-wide model for the normal eukaryotic DNA replication fork. *Proc. Natl Acad. Sci. USA* **107**, 17674–17679 (2010).
35. Pursell, Z. F., Isoz, I., Lundstrom, E. B., Johansson, E. & Kunkel, T. A. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**, 127–130 (2007).
36. Raghuraman, M. K. *et al.* Replication dynamics of the yeast genome. *Science* **294**, 115–121 (2001).
37. Nick McElhinny, S. A. *et al.* Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl Acad. Sci. USA* **107**, 4949–4954 (2010).
38. Kao, H. I. & Bambara, R. A. The protein components and mechanism of eukaryotic Okazaki fragment maturation. *Crit. Rev. Biochem. Mol. Biol.* **38**, 433–452 (2003).
39. Chon, H. *et al.* RNase H2 roles in genome integrity revealed by unlinking its activities. *Nucleic Acids Res.* **41**, 3130–3143 (2013).
40. Boyle, A. P. *et al.* High-resolution genome-wide *in vivo* footprinting of diverse transcription factors in human cells. *Genome Res.* **21**, 456–464 (2011).
41. Cooper, G. M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* **15**, 901–913 (2005).
42. Schmidt, S. *et al.* Hypermutable non-synonymous sites are under stronger negative selection. *PLoS Genet.* **4**, e1000281 (2008).
43. Vengrova, S. & Dalgaard, J. Z. The wild-type *Schizosaccharomyces pombe* *mat1* imprint consists of two ribonucleotides. *EMBO Rep.* **7**, 59–65 (2006).
44. Ghodgaonkar, M. M. *et al.* Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol. Cell* **50**, 323–332 (2013).
45. Liberti, S. E., Larrea, A. A. & Kunkel, T. A. Exonuclease 1 preferentially repairs mismatches generated by DNA polymerase α . *DNA Repair* **12**, 92–96 (2013).
46. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nature Struct. Mol. Biol.* **20**, 14–22 (2013).
47. Villar, D., Flicek, P. & Odom, D. T. Evolution of transcription factor binding in metazoans - mechanisms and functional implications. *Nature Rev. Genet.* **15**, 221–233 (2014).
48. Clausen, A. R. *et al.* Tracking replication enzymology *in vivo* by genome-wide mapping of ribonucleotide incorporation. *Nature Struct. Mol. Biol.* <http://dx.doi.org/10.1038/nsmb.2957> (2015).
49. Koh, K. D., Balachander, S., Hesselberth, J. R. & Storici, F. Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. *Nature Methods* <http://dx.doi.org/10.1038/nmeth.3259> (2015).
50. Daigaku, Y. *et al.* A global profile of replicative polymerase usage. *Nature Struct. Mol. Biol.* (in the press).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank N. Hastie and P. Burgers for discussions, I. Adams, J. Caceres, T. Aitman and P. Heyn for comments on the manuscript, and A. Gallacher for technical assistance. We are indebted to J. Williams, A. Clausen and T. Kunkel for sharing yeast strains and unpublished data, and to S. Cerritelli and R. Crouch for *RNH201* vectors. Funding: MRC Centenary Award to M.A.M.R.; MRC and Lister Institute of Preventive Medicine to A.P.J.; MRC and Medical Research Foundation to M.S.T.

Author Contributions M.A.M.R. and J.D. performed experiments. H.K., S.M.d.P. and M.S.T. performed computational analysis. M.S.T., M.A.M.R. and A.P.J. designed the study, developed emRiboSeq and wrote the manuscript.

Author Information Next-generation sequencing data submitted to the Gene Expression Omnibus (GEO) under accession number GSE64521; R and PERL code submitted to GitHub repository taylorLab/LaggingStrand (<https://github.com/taylorLab/LaggingStrand>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.P.J. (andrew.jackson@igmm.ed.ac.uk) or M.S.T. (martin.taylor@igmm.ed.ac.uk).

METHODS

Yeast reference genome and annotation. All analyses were performed on the sacCer3 (V64) *S. cerevisiae* reference genome assembly. Data sets originally obtained with coordinates on other assemblies were projected into the sacCer3 assembly using liftOver (v261)⁵¹ with the corresponding chain files obtained from <http://www.yeastgenome.org>. All regions of the sacCer3 genome were used for read alignment but analyses including strand ratios and all rate estimates excluded the following multi-copy regions: the mitochondrial genome, rDNA locus chrXII:459153–461153 and any 100-nucleotide segment with mappability score of <0.9 (gem-mappability⁵² with k -mer = 100). In total, this masked 951,532 nucleotides (7.8%) of the reference genome. Gene structure annotations were the *Saccharomyces* Genome Database (SGD) consensus annotations extracted from the University of California, Santa Cruz (UCSC) genome browser in November 2013. Annotated origins of replication were obtained from ref. 53. DNase I hypersensitive sites and footprints were obtained from ref. 54, and nucleosome position, occupancy and positional fuzziness (positional heterogeneity) measures were from ref. 55. Yeast replication timing data was obtained from ref. 36, where we have plotted the percentage of heavy–light (replicated) DNA (pooled samples data set). Higher percentage indicates earlier average replication time.

Yeast polymorphisms and between species substitution rates. Yeast polymorphism data was obtained from the *Saccharomyces* Genome Resequencing project⁵⁶. A polymorphic difference between any of the 37 sequenced *S. cerevisiae* strains was called as a polymorphic site. Sites with $n > 2$ alleles were only counted once as a polymorphic site. Only nucleotide point substitutions were considered, insertions and deletions were excluded. The polymorphism rate reported is the number of polymorphic sites divided by the number of sacCer3 sites with sequence coverage in at least two additionally sequenced strains.

Yeast between-species substitution rates were calculated from MultiZ stacked pairwise alignments obtained from the UCSC genome browser (Supplementary Table 1). Alignments for five *sensu stricto* yeast species (*S. cerevisiae*, *S. paradoxus*, *S. mikatae*, *S. kudriavzevii* and *S. bayanus*) were extracted from the original seven species alignment. The reference assembly names and phylogenetic relationship are represented by the tree (((sacCer3, sacPar), sacMik), sacKud, sacBay). Substitution rates were calculated over whole chromosomes using baseml from the paml⁵⁷ package (version 4.6) under the HKY85 substitution model with $\text{ncatG} = 5$ categorical gamma. Per-nucleotide relative rate estimates (branch length multipliers) were obtained over the sacCer3 genome.

Human conservation measures. GERP scores⁴¹ were used as a measure of between species nucleotide diversity across 46 vertebrate species. Single-nucleotide resolution bigWig files were obtained from UCSC genome browser (hg19). For consistency of presentation with plots of polymorphism rate and yeast between-species nucleotide substitution rate, the y axes in plots showing GERP scores have been inverted so that greater constraint is low and greater diversity is high.

OF sequence processing. OF sequence data was obtained from ref. 17 (GEO accession GSM835651). Analysis primarily focused on the larger ‘replicate’ library but results were confirmed in the ‘sample’ library (GEO accession GSM835650). The OF strand ratio was calculated as the sum of per nucleotide read coverage on the forward strand divided by the same measure for reverse strand reads. OF strand ratios were calculated in windows of 2,001 nucleotides. A pseudo count of 1 read-covered nucleotide was added to both strands in each window to avoid divisions by zero. Results shown are for de-duplicated read data (identical start and end coordinates were considered duplicates). De-duplication minimises potential biases in PCR amplification, qualitatively similar results were obtained with non-de-duplicated data and support identical conclusions.

Rather than using separate Okazaki 5′ and 3′ end counts that did not always correlate well, probably due to amplification biases, sequencing and size selection biases; we produced a normalized OJ rate measure. This is the average of (1) the fraction of upstream OFs that terminate with a 3′ end at a focal nucleotide, and (2) the fraction of downstream OFs whose 5′ end is at the focal nucleotide. The upstream and downstream coverage measures were based on mean Okazaki read coverage for the nucleotides located between 5 and 12 nucleotides upstream (downstream) of the focal 3′ (5′) end. This OJ rate was calculated at single nucleotide resolution over both strands of the sacCer3 genome.

EmRiboSeq alignment and processing. Sequence reads (see Supplementary Table 2 for runs and read numbers) were aligned to the unmasked sacCer3 genome with bowtie2 (version 2.0.0). Subsequent filtering and format conversion were performed using Samtools (version 0.1.18) and BEDTools (version 2.16.2). Only reads with a mapping quality score >30 were kept for analysis. As there had been no pre-sequencing amplification, de-duplication was not performed. Read 5′-end counts were summed per strand at single nucleotide resolution over the yeast reference genome. Note that under the emRiboSeq protocol, the ribonucleotide incorporation site would be one nucleotide upstream and on the opposite strand to the mapped read 5′ end. To facilitate comparison between libraries of

differing read depth, read counts were normalized to sequence tags per million mapped into the non-masked portion of the genome.

Defining TF binding sites. Reb1 and Rap1 ChIP-exo data was obtained from ref. 58 (Sequence Read Archive accession SRA044886). Sequence bar codes were clipped and sequences sorted using Perl (version 5.18.2). Reads were aligned using bowtie2 (version 2.0.0). Following the previously published protocol⁵⁸ up to three mismatches across the length of each tag sequence were allowed, and the 3′ most 6 base pairs (bp) removed. Peaks were called with MACS (version 2.0.10). Following ref. 58, sites were defined as monomer if no other peaks were present within 100 bp. Where two or more peaks were present within 100 bp the peak with the highest occupancy was labelled as the primary peak. Telomeric sites were excluded using annotations within the sacCer3 sgdOther UCSC table (<http://www.yeastgenome.org>). The presence or absence of a motif was determined using the Motif Occurrence Detection Suite (MOODS)⁵⁹ (version 1.0.1). Consensus binding motifs positional weight matrices were obtained from JASPAR⁶⁰ (<http://jaspar.genereg.net/>). The matching motif significance threshold was set at 0.005. Multiple peaks were aligned ($x = 0$) to the midpoint of the JASPAR defined motif. Human TF binding sites were defined using ChIP-seq data (Supplementary Table 1) as for yeast, except that the peak clustering threshold was reduced to 50 nucleotides.

Computational and statistical analyses. Analysis and all statistical calculations were performed in R (version 3.0.0). Lines of fit used the smooth.spline function with degrees of freedom: Fig. 1a, 18 degrees; Fig. 1b, 34 degrees; Fig. 3d–f, 80 degrees of freedom (strand ratio calculated in 2,001-nucleotide consecutive windows). Sliding window averages used the rollapply function from the Zoo package with centre alignment and null padding. Pearson’s correlation was performed with the cor.test function in R, paired Student’s t -test with the t.test function, Mann–Whitney tests with the wilcox.test function and lowess (locally weighted scatterplot smoothing) with the lowess function and default parameters.

No statistical methods were used to predetermine sample size.

Rate estimates with compositional correction. Polymorphism and OJ rates were calculated separately for each nucleotide (A, T, C or G) and the average of these for rates used as the reported or plotted measure for a nucleotide site or group of sites. This corrects for mononucleotide compositional biases that are abundant when sampling specific features of a genome. The between-species relative substitution rate calculation incorporates a compositional correction. The rate estimates shown are the number of observations divided by the number of sites with non-missing data.

Trinucleotide preserving shuffles. Every nucleotide of the sacCer3 genome was assigned to one of 64 categories based on the identity of that nucleotide and its flanking nucleotides. A vector of transformations was produced by swapping the genomic coordinate of a nucleotide for one with an identical category chosen at random. Swaps between masked and unmasked sites (see above) were prevented. 100 such vectors were produced. For a set of stacked coordinates (for example, Fig. 1a comprising 27,586 sequences, each of 251 nucleotides), every nucleotide of every sequence was substituted through the transformation vector, for a randomly selected proxy, matched for the same trinucleotide context and their corresponding rate or annotation used. This provides a compositionally well-matched null expectation. With 100 independent transformation vectors we provide empirically derived 95% confidence bounds and standard deviations on those null expectations. For human sites, shuffles were confined to sequences flanking the region of interest (100–300 nucleotides distant from the binding site for TF analysis and 1,000–2,000 nucleotides distant for DNase I footprint analysis). Human genomic coordinates in the ENCODE ‘Duke Excluded Regions’⁶¹ and those positions with a uniqueness score of <0.9 (gem-mappability⁵² with k -mer = 100) were excluded from shuffles.

Sites selected for analysis. Thresholds were applied to define specific subsets of sites to be evaluated. For the presented data (Fig. 1a) nucleosomes with an occupancy of >80%, positional fuzziness⁵⁵ of <30, with at least 30 OF reads over them, and located more than 200 nucleotides from transcription start sites were used. Other combinations (Extended Data Fig. 1) of these parameters gave qualitatively similar results and support the same conclusions. Reb1 (and Rap1) sites were defined as the primary ChIP-exo peak at a site, with sequences aligned ($x = 0$) to the centre of the highest scoring Reb1/Rap1 position weight matrix match within 50 nucleotides of the ChIP-exo peak summit. DNase I footprints from 41 human cell types were previously combined⁶² into consensus footprints (combined.fps.gz). We intersected the combined footprints with those found in each cell type using BEDtools (version 2.17.0) to identify the subset ($n = 33,530$) that were detected in all 41 cell types. The left-edge coordinate as defined in the combined footprint file was used as the focal nucleotide ($x = 0$) for analysis.

Comparison of polymorphism rates. The five nucleotide positions downstream and the five upstream of the focal OJ position (excluding $x = 0$ in both cases) were scored for their polymorphism rate (Fig. 2b). Rate deltas were calculated as upstream minus downstream in 100 bootstrap replicates and a paired two-sided t -test

performed against the same calculation performed on 100 trinucleotide preserving genome shuffles of the same sites. This tests whether the difference in rate between upstream and downstream positions is greater in the observed data than the shuffled data.

DNA purification. Yeast strains were grown at 30 °C in YPDA to mid-log phase (see Supplementary Table 3 for a list of strains) or to saturation for stationary phase. Per 5 $A_{600\text{ nm}}$ units, cell pellets were resuspended in 200 μl lysis buffer (2% Triton X-100, 1% SDS, 0.5 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA). An equal volume of TE-equilibrated phenol and glass beads (0.40–0.60 mm diameter, Sartorius) were added, and cells lysed by vortexing for 2 min; 200 μl TE buffer was then added, followed by an additional 1 min of vortexing. After centrifugation, the aqueous phase was further extracted with equal volumes of phenol:chloroform:isoamylalcohol (25:24:1) and chloroform. Total nucleic acids were precipitated with 1 ml of 100% ethanol, and dissolved in 0.5 M NaCl. RNA was degraded by treatment with 10 μg RNase A (Roche) for 1 h at room temperature. DNA was finally purified with an equal volume of Ampure XP beads (Beckman Coulter) and eluted in nuclease-free water. For library preparations DNA was isolated from up to 40 $A_{600\text{ nm}}$ units.

Alkaline gel electrophoresis. Isolated genomic DNA (0.5 μg) was treated with recombinant RNase H2, purified as previously described⁶³ and ethanol precipitated. DNA pellets were dissolved in alkaline loading dye and separated on 0.7% agarose gels (50 mM NaOH, 1 mM EDTA) as previously described²⁷, and stained with SYBR Gold (Life Technologies). Densitometry measurements and derivation of ribonucleotide incorporation rates as previously described²⁸. Percentage genome contribution for each replicative polymerase (x) was calculated using the following formula: $N_{\Delta\text{polx}} \cdot F_{\text{polx}} / (N_{\text{pola}} \cdot F_{\text{pola}} + N_{\text{pold}} \cdot F_{\text{pold}} + N_{\text{pole}} \cdot F_{\text{pole}})$, with $N_{\Delta\text{polx}}$ the number of ribonucleotides incorporated in one yeast genome for the mutant polymerase, above that detected in the Δrnh201 POL strain, measured on the same alkaline gel, and F_{polx} the frequency of incorporation by that polymerase (see Fig. 3a).

EmRiboSeq library preparation and sequencing. DNA was sonicated using a Bioruptor Plus (Diagenode) to achieve an average fragment length of approximately 400 bp. Fragmented DNA was concentrated by ethanol precipitation and size selected using 1.2 volumes of Ampure XP. DNA was quantified by nanodrop (Thermo Scientific) and up to 5 μg was used for NEBNext End Repair and dA-Tailing (New England Biolabs) following the manufacturer's guidelines. After the end-repair reaction, DNA was purified using 1.2 volumes of Ampure XP. Subsequent steps were performed in the presence of Ampure XP beads, capturing the DNA by adding NaCl and PEG_{8,000} to final concentrations of 1.25 M and 10%, respectively. The trP1 adaptor (see below) was attached using NEBNext Quick Ligation with 120 pmol of adaptor per microgram of DNA for 14–18 h at 16 °C.

Terminal transferase (NEB) was then used to block any free 3' ends with ddATP for 2 h at 37 °C, with 20 U of TdT per microgram of DNA. After Ampure XP purification, beads were removed and DNA nicked using recombinant RNase H2 (10 pmol μg^{-1} of library) or Nb.BtsI (NEB; 10 U μg^{-1}) for 2 h at 37 °C. RNase H2 purification and reaction conditions were as previously described⁶³. Enzymes were inactivated by heating at 80 °C for 20 min, and DNA was purified using 1.8 volumes of Ampure XP. Shrimp alkaline phosphatase (Affymetrix; 5 U) was then used to remove 5' phosphates at 37 °C (1 h per μg of library). After heat inactivation for 15 min at 65 °C and Ampure XP purification, DNA was denatured by heating at 95 °C for 5 min and snap cooling. Subsequently, A adaptor (see below; 120 pmol μg^{-1} of library) was attached using NEBNext Quick Ligation for 14–18 h at 16 °C. Fragments

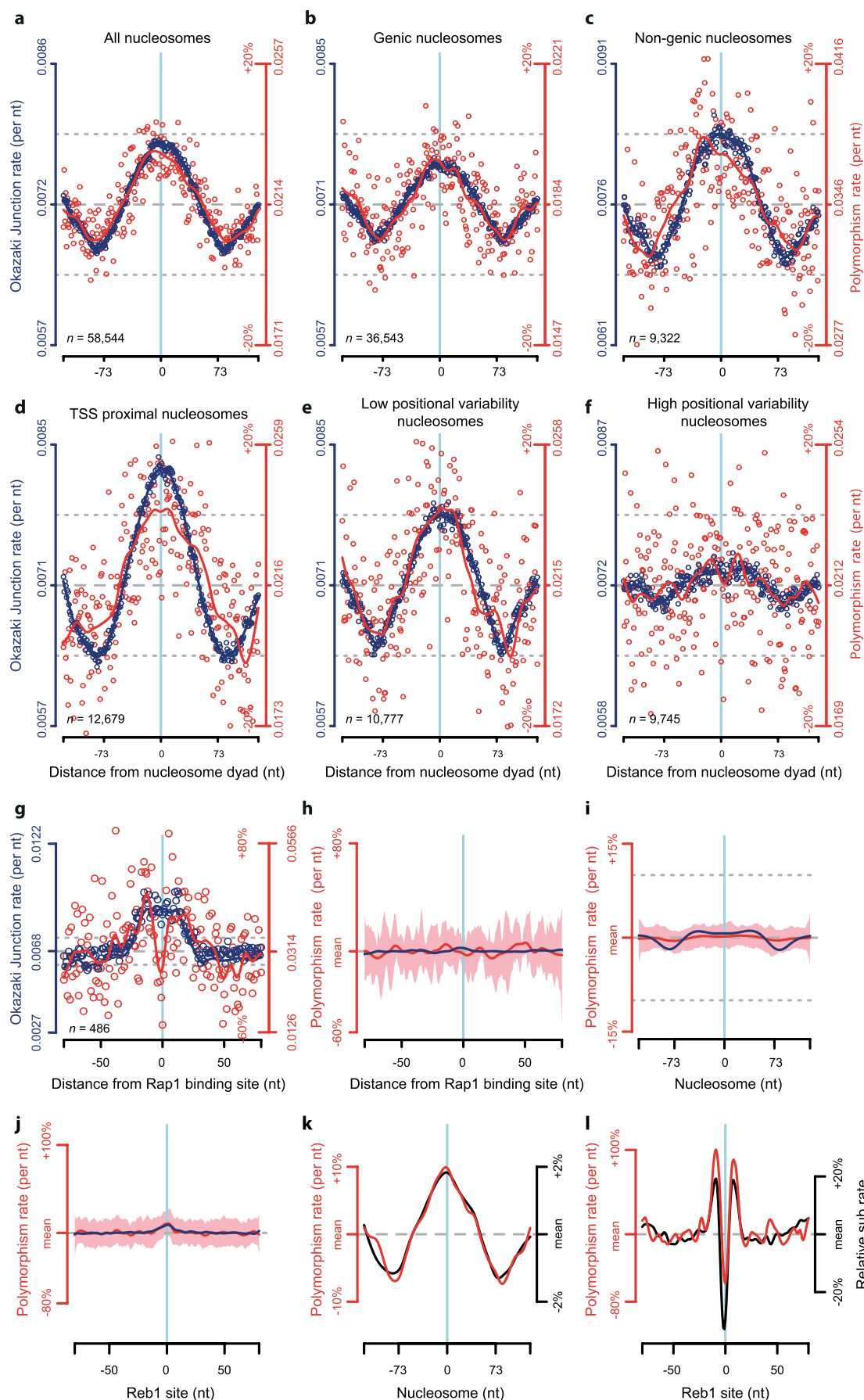
with biotinylated A adaptor were captured on streptavidin-coupled M-280 Dynabeads (Life Technologies) following the manufacturer's guidelines, and non-biotinylated strands were released in 0.15 M NaOH. Single-stranded fragments were concentrated by ethanol precipitation.

Phusion Flash High-Fidelity PCR Master Mix (Thermo Scientific) was then used for second strand synthesis with primer A to produce a double stranded library. Size selection of fragments between 200 and 300 bp in size was performed using 2% E-Gel EX (Life Technologies). Finally, this library was quality checked and quantified using a 2100 Bioanalyzer (Agilent Technologies) before emulsion PCR, using the Ion Torrent One Touch, and next generation sequencing on the Ion Torrent PGM or Proton platform (Life Technologies).

Oligonucleotides and adaptor design. Custom oligonucleotides were synthesized by Eurogentec. Adaptor primer pairs were annealed by heating at 95 °C for 5 min and cooling gradually. Sequences of the adaptor primer pairs were as follows. Adaptor 1 (trP1): trP1-top, 5'-CCTCTCTATGGCAGTCCGGTGTAT-phosphorothioate-T-3'; trP1-bottom, 5'-phosphate-ATCACCGACTGCCCATAGAGAGGC-dideoxy-3'. Adaptor 2 (A): A-top, 5'-phosphate-CTGAGTCGGA GACACGCAGGGATGAGATGG-dideoxy-3'; A-bottom, 5'-biotin-CCATCTC ATCCCTCGGTGTCTCCGACTCAGNNNNNNN-C3 phosphoramidite-3'. The sequence for primer A used in second strand synthesis was 5'-CCATCTCATCCTGCGTGTCTCCGAC-3'.

Data sources, sequencing data and *S. cerevisiae* strains. Documented in Supplementary Tables 1–3.

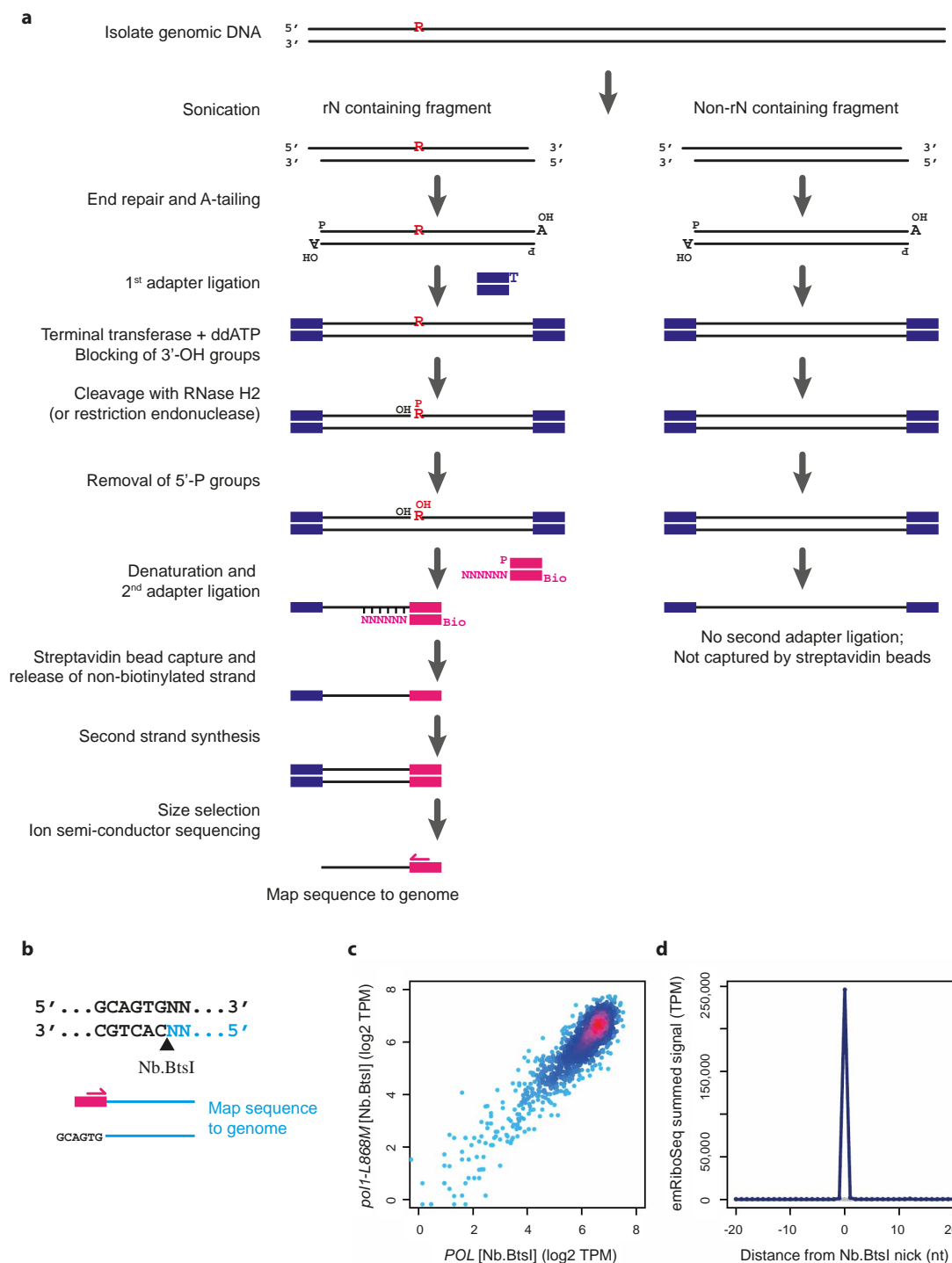
51. Kuhn, R. M., Haussler, D. & Kent, W. J. The UCSC genome browser and associated tools. *Brief. Bioinform.* **14**, 144–161 (2013).
52. Derrien, T. *et al.* Fast computation and applications of genome mappability. *PLoS ONE* **7**, e30377 (2012).
53. Eaton, M. L., Galani, K., Kang, S., Bell, S. P. & MacAlpine, D. M. Conserved nucleosome positioning defines replication origins. *Genes Dev.* **24**, 748–753 (2010).
54. Hesselberth, J. R. *et al.* Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. *Nature Methods* **6**, 283–289 (2009).
55. Jiang, C. & Pugh, B. F. A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genome. *Genome Biol.* **10**, R109 (2009).
56. Liti, G. *et al.* Population genomics of domestic and wild yeasts. *Nature* **458**, 337–341 (2009).
57. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
58. Rhee, H. S. & Pugh, B. F. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**, 1408–1419 (2011).
59. Korhonen, J., Martinmaki, P., Pizzi, C., Rastas, P. & Ukkonen, E. MOODS: fast search for position weight matrix matches in DNA sequences. *Bioinformatics* **25**, 3181–3182 (2009).
60. Mathelier, A. *et al.* JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **42**, D142–D147 (2014).
61. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
62. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).
63. Reijns, M. A. *et al.* The structure of the human RNase H2 complex defines key interaction interfaces relevant to enzyme function and human disease. *J. Biol. Chem.* **286**, 10530–10539 (2011).



Extended Data Figure 1 | Increased OJ and polymorphism rates correlate at binding sites of different nucleosome classes and at Rap1 binding sites.

a–f, OJ and polymorphism rates are strongly correlated for different classes of nucleosomes. Data presented as in Fig. 1a, for different sub-classes of *S. cerevisiae* nucleosomes, demonstrating that OJ and polymorphism rates covary in all cases. Transcription start site proximal nucleosomes (**d**) are probably subject to strong and asymmetrically distributed selective constraints, which is likely to explain the modestly reduced correlation for this subset. Such transcription start site proximal nucleosomes were excluded from analyses of other categories presented (**b**, **c**, **e**, **f**), except ‘all nucleosomes’ (**a**). **g**, OJ and polymorphism rates are correlated for the *S. cerevisiae* TF, Rap1. Data presented, as for Reb1 in Fig. 1b, show increased OJ and polymorphism rates

around its binding site, with a dip corresponding to its central recognition sequence. **h–j**, Increased polymorphism and OJ rates at Rap1 (**h**), nucleosome (**i**) and Reb1 (**j**) binding sites are not caused by biases in nucleotide content. Distributions calculated as for **g**, Fig. 1a and b, respectively, using a trinucleotide preserving genome shuffle. Pink shaded areas denote 95% confidence intervals for nucleotide substitution rates (100 shuffles). **k**, **l**, Polymorphism (red) and between-species (black) substitution rates are highly correlated for nucleosome (**k**) and Reb1 (**l**) binding sites. Best fit splines shown only. *y* axes scaled to demonstrate similar shape distribution. Values plotted as percentage relative to the mean rate for all data points (central 11 nucleotides excluded for calculation of mean in **g**, **l**).

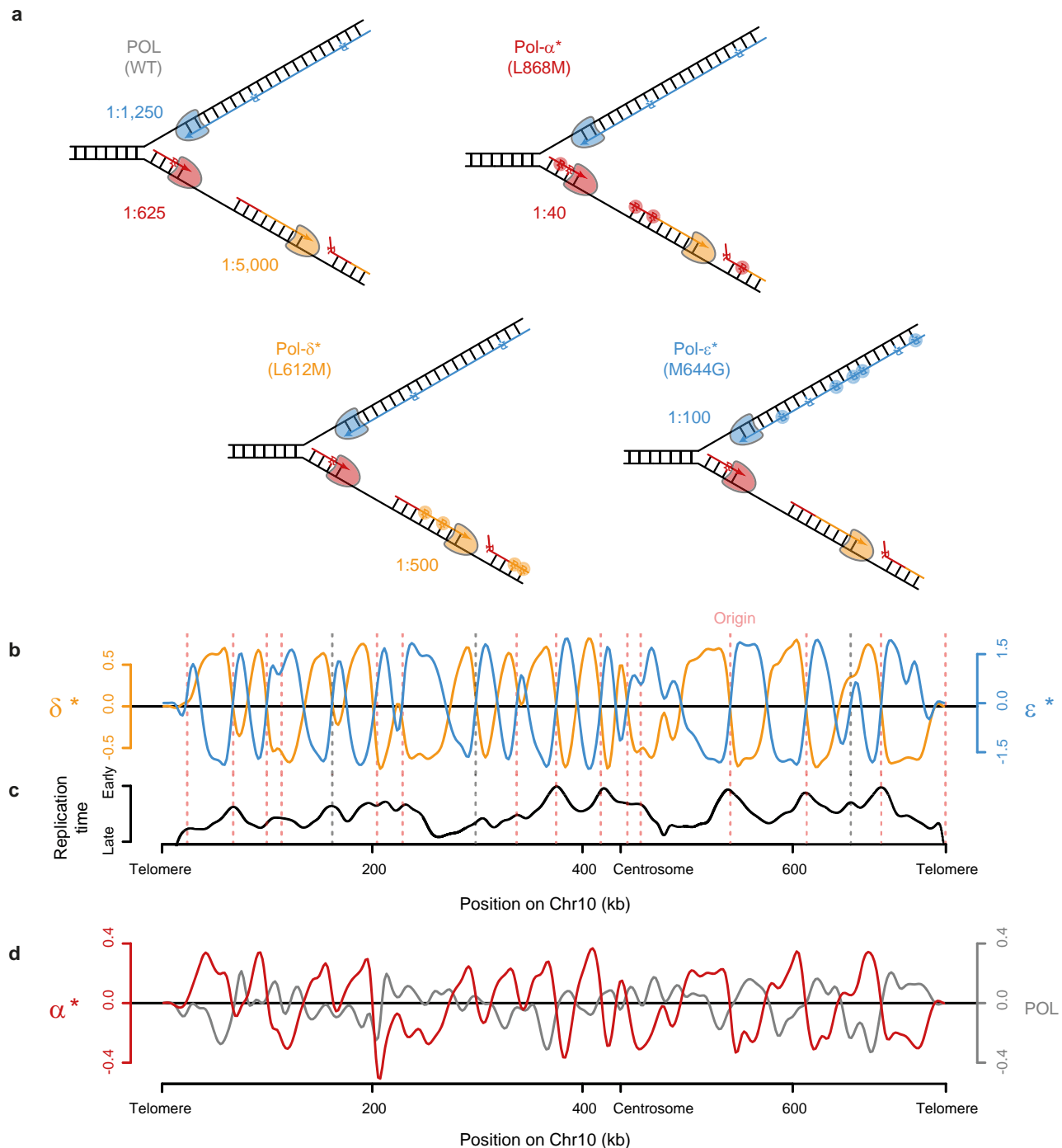


Extended Data Figure 2 | EmRiboSeq methodology and validation.

a, Schematic of emRiboSeq library preparation. rN, ribonucleotide.

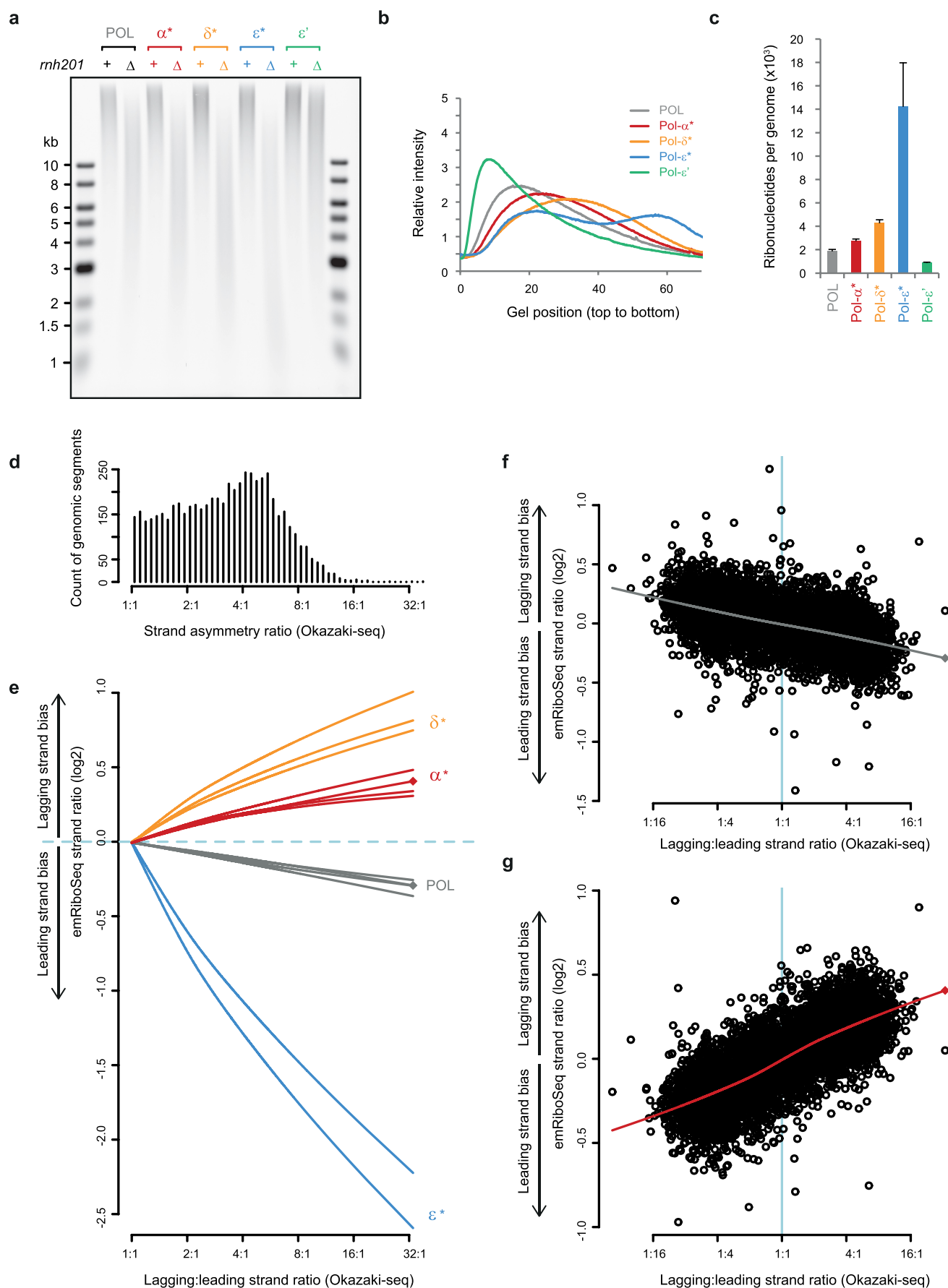
b–d, Validation of strand-specific detection of enzymatically generated nicks through linker-ligation. Nb.BtsI nicking endonuclease cleaves the bottom strand of its recognition site releasing a 5' fragment (cyan) with a free 3'-OH group after denaturation, to which the sequencing adaptor (pink) is ligated, allowing sequencing and mapping of this site to the genome (**b**). Nb.BtsI libraries have high reproducibility between *Δrnh201* POL and *Δrnh201* Pol-α* (*pol1-L868M*) strains after normalizing read counts to sequence tags per million

(TPM). Bona fide Nb.BtsI sites were equally represented, at maximal frequency, in both libraries (**c**). Those with lower frequencies represented sites in close proximity to other Nb.BtsI sites, causing their partial loss during size selection. Additionally, Nb.BtsI-like sites were detected as the result of star activity. Libraries were also prepared using BciVI restriction enzyme digestion, that did not show such star activity (data not shown), allowing calculation of the site specificity for the method (>99.9%). Summed signal at Nb.BtsI sites shows >99.9% strand specificity (blue, correct strand; grey, opposite strand) and >99% single nucleotide resolution (**d**).



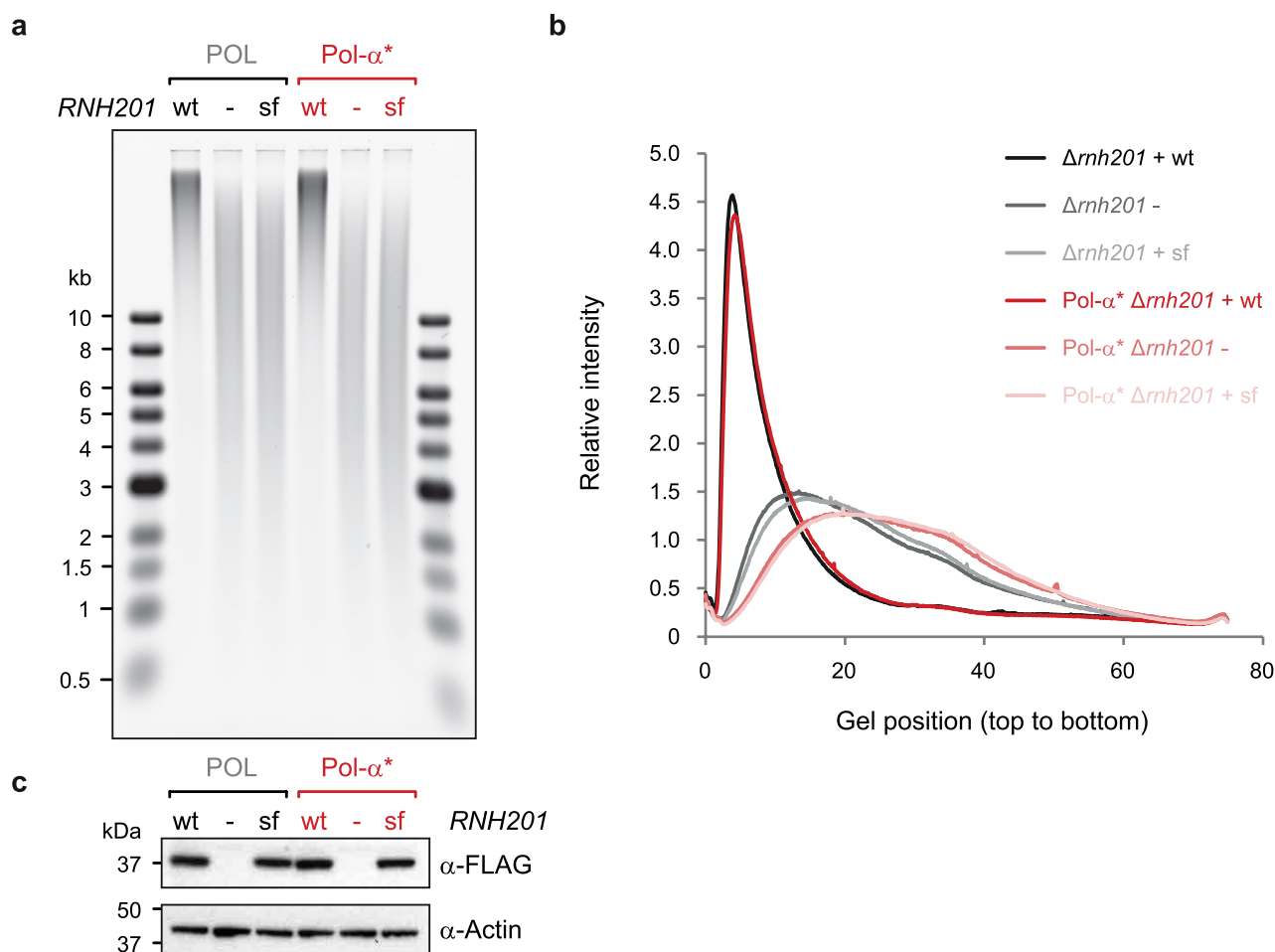
Extended Data Figure 3 | Mapping replicative polymerase DNA synthesis using emRiboSeq. **a**, Point mutations in replicative polymerases elevate ribonucleotide incorporation rates, permitting their contribution to genome synthesis to be tracked. Schematic of replication fork with polymerases and their ribonucleotide incorporation rates (refs 27, 30 and J. S. Williams, A. R. Clausen & T. A. Kunkel, personal communication) as indicated (POL denotes wild-type polymerases; asterisk denotes point mutants). Embedded ribonucleotides indicated by 'R'; additional incorporation events due to polymerase mutations highlighted by shaded circles. **b**, **c**, Mapping of leading/

lagging-strand synthesis by Pol- δ^* and Pol- ϵ^* yeast strain using emRiboSeq (as in Fig. 3) highlights both experimentally validated (pink dotted lines) and putative (grey dotted lines) replication origins. These often correspond to regions of early replicating DNA³⁶ (**c**). **d**, Pol- α^* DNA is detected genome-wide by emRiboSeq as a component of the lagging strand in stationary phase yeast, as shown by the opposite pattern for a polymerase wild-type strain. Strand ratios are shown as best-fit splines with 80 degrees of freedom, y axes show \log_2 of the strand ratio calculated in 2,001-nucleotide windows (**b–d**).



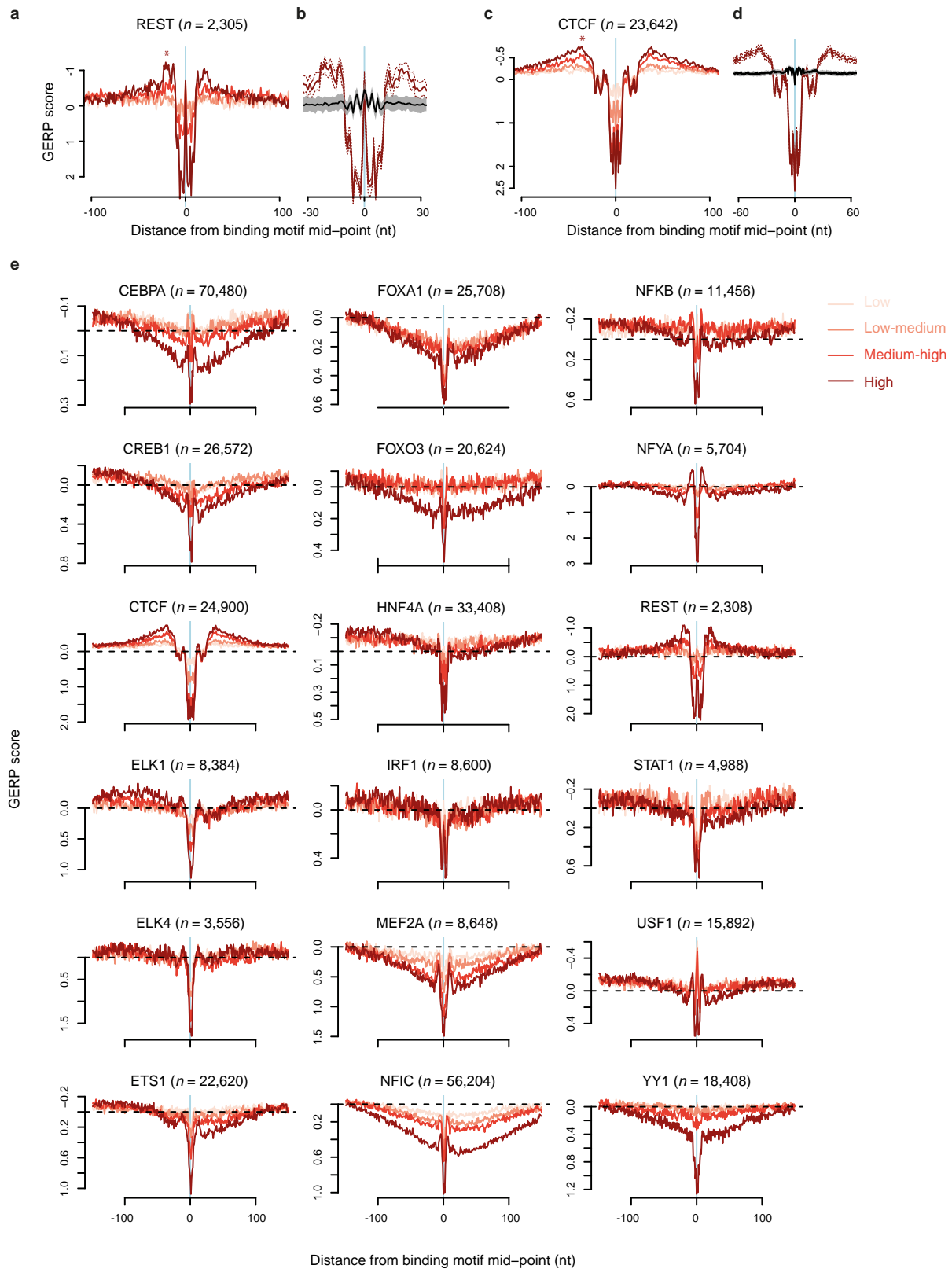
Extended Data Figure 4 | Quantification of *in vivo* ribonucleotide incorporation by replicative polymerases. **a, b,** Representative alkaline gel electrophoresis of genomic DNA from yeast strains with mutant replicative DNA polymerases (**a**), with accompanying densitometry plots (**b**). Embedded ribonucleotides are detected by increased fragmentation of genomic DNA following alkaline treatment in an RNase H2-deficient ($\Delta rnh201$) background. Increased rates are seen with all three mutant polymerases (indicated by asterisk, as defined in Extended Data Fig. 3a), and are reduced in Pol- ϵ' which contains the point mutation Met644Leu, a mutation that increases selectivity for dNTPs over rNTPs²⁷. **c,** Quantification of average ribonucleotide incorporation in polymerase mutants from four independent experiments. DNA isolated from mid-log phase cultures; error bars denote s.e.m. Overall ribonucleotide content is the product of incorporation frequency and the total contribution of each polymerase, resulting in the total ribonucleotide content detected to be highest for Pol- ϵ^* (14,200 per genome), followed by Pol- δ^* (4,300 per genome), Pol- α^* (2,700 per genome), POL (1,900 per genome) and Pol- ϵ' (860 per genome). **d,** Most of the yeast genome exhibits directional asymmetry in replication (median 4:1 strand ratio). Count of genomic

segments calculated for consecutive 2,001-nucleotide windows over the yeast genome based on reanalysis of OF sequencing data¹⁷ denoted as 'Okazaki-seq'. The strand asymmetry ratio was calculated after re-orienting all regions such that the predominant lagging strand was the forward strand. **e–g,** Genome-wide quantification of strand-specific incorporation of wild-type and mutant replicative DNA polymerases determined by emRiboSeq reflects their roles in leading- and lagging-strand replication. A close to linear correlation with Okazaki-seq strand ratios is observed. The strand ratio preference for lagging-strand ribonucleotide incorporation for independent libraries (including stationary phase libraries for POL and Pol- α^* , marked by diamonds) was plotted against the lagging:leading-strand ratio determined using Okazaki-seq data (only ratios $\geq 1:1$ for the latter are shown for clarity). There was high reproducibility between experiments in strand ratio preferences. Lines are lowess smoothed (see Methods) representations of the full data sets (representative examples given in **f** and **g**). **f, g,** Scatter plots illustrating the individual strand ratio data points for 2,001-nucleotide windows, for stationary phase POL (**f**) and Pol- α^* (**g**) yeast. Pearson's correlation = 0.49, $P < 2.2 \times 10^{-16}$ for POL (**f**); correlation = 0.75, $P < 2.2 \times 10^{-16}$ for Pol- α^* (**g**).



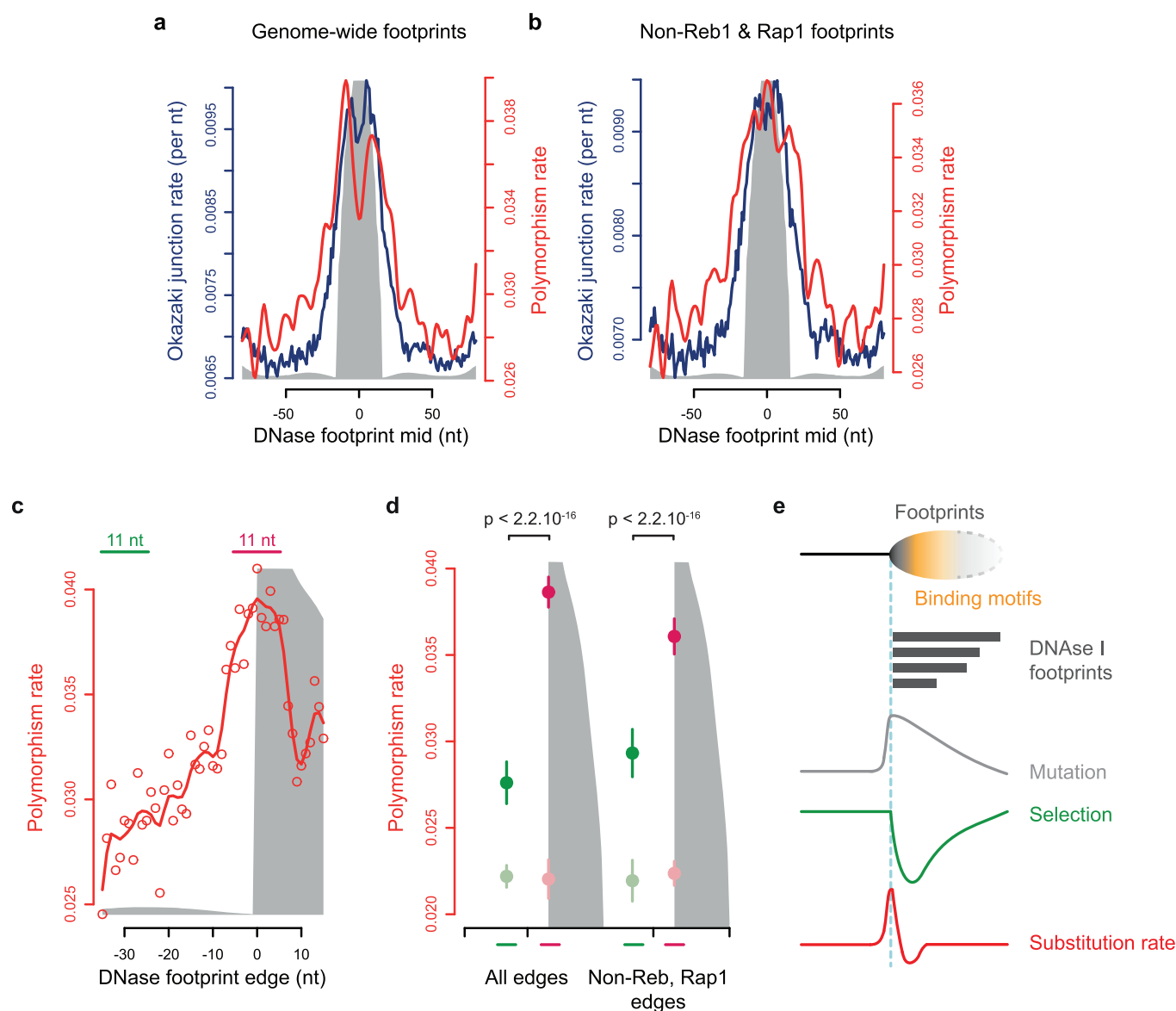
Extended Data Figure 5 | Pol- α -synthesized DNA retention is independent of RNase H2 processing of RNA primers. **a, b,** The ribonucleotide content of genomic DNA is unchanged between $\Delta rnH201$ strains transformed with empty vector (–) or vector expressing Rnh201 separation-of-function mutant (sf), that retains the ability to cleave RNA:DNA hybrids, including RNA primers, but cannot cleave single embedded ribonucleotides³⁹. In contrast, the same vector expressing wild-type Rnh201 (wt) fully rescues alkaline sensitivity of the

DNA. As complementation with the separation-of-function mutant had no detectable effect on the ribonucleotide content seen in the Pol- α (Leu868Met) $\Delta rnH201$ strain, retention of Pol- α -synthesized DNA appears to be independent of a putative role for RNase H2 in RNA primer removal. Representative result shown for $n = 3$ independent experiments. **c,** Wild-type and mutant Rnh201 are expressed at equal levels, as shown by immunodetection of the C-terminal FLAG tag. Loading control, actin.



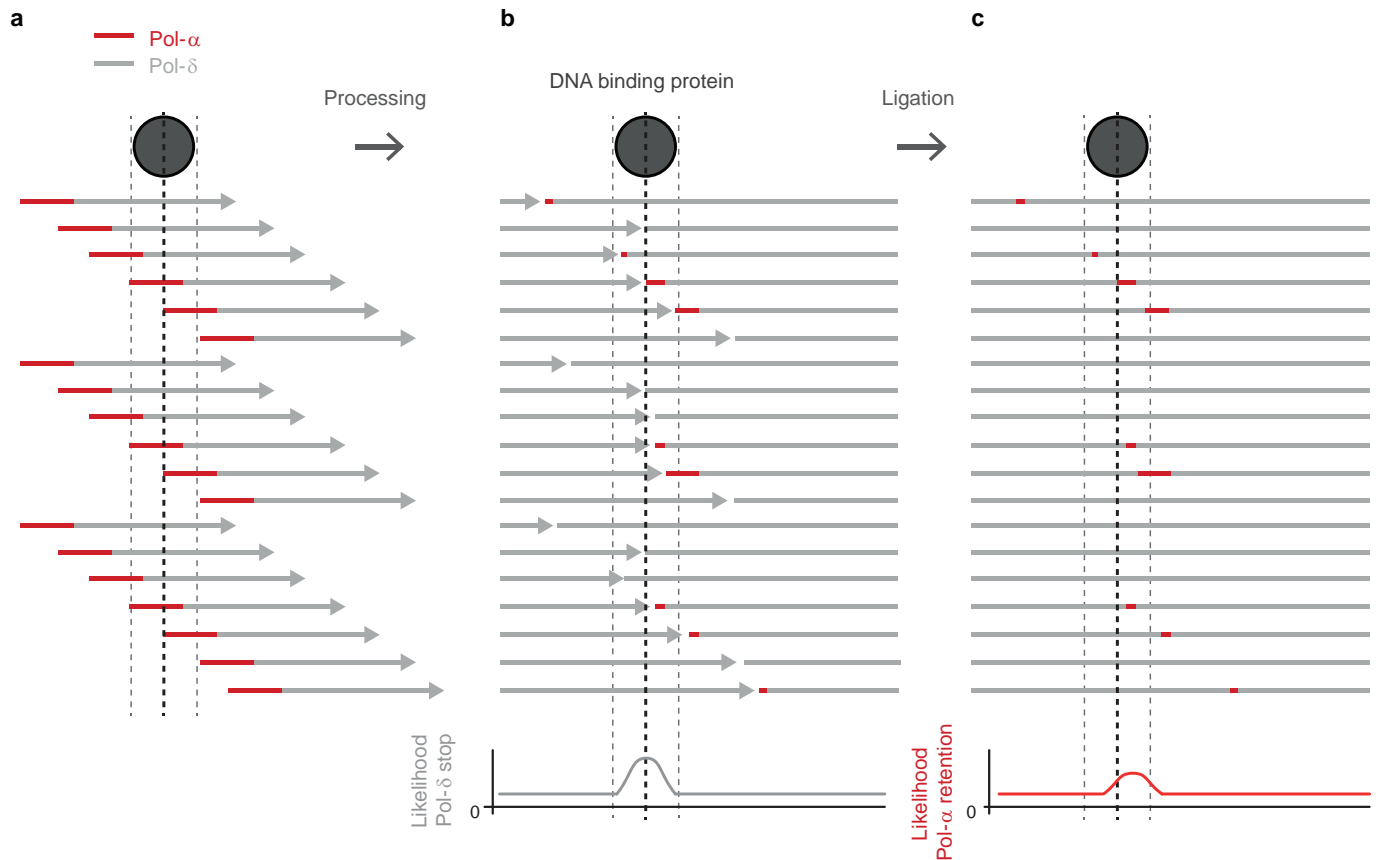
Extended Data Figure 6 | Elevated substitution rates are observed adjacent to many human TF binding sites. **a–d**, Nucleotide substitution rates (plotted as GERP scores) are elevated immediately adjacent to REST (**a**, **b**) and CTCF binding sites (**c**, **d**). Colour intensity shows quartiles of ChIP-seq peak height (pink to brown: lower to higher), reflecting strength of binding/occupancy. Stronger binding correlates with greater increases of proximal substitution rate in the 'shoulder' region (asterisk). Increased substitution rates are not a consequence of local sequence composition effects (**b**, **d**). Strongest binding

quartile of sites (brown) is shown compared to a trinucleotide preserving shuffle (black) based on the flanking sequence (100–300 nucleotides from motif midpoint) of the same genomic locations. Brown dashed line and grey shading denote 95% confidence intervals. **e**, Substitution rates plotted as GERP scores for human TF binding sites identified in ChIP-seq data sets (in conjunction with binding site motif). Sites aligned ($x = 0$) on the midpoint of the TF binding site within the ChIP-seq peak (colours as for **a–d**). Dashed black line shows $y = 0$, the genome wide expectation for neutral evolution.



Extended Data Figure 7 | OJ and polymorphism rates are increased at yeast DNase I footprints. **a, b**, DNase I footprint edges correspond, genome-wide, to increased OJ rates and locally elevated polymorphism rates in *S. cerevisiae* (**a**), a pattern that is maintained when footprints associated with Reb1 and Rap1 binding sites are excluded (**b**). Genome-wide DNase I footprints ($n = 6,063$) and excluding those within 50 nucleotides of a Reb1 or Rap1 binding site ($n = 5,136$) were aligned to their midpoint. **c, d**, Aligning DNase I footprints on their left edge rather than midpoint (to compensate for substantial heterogeneity in footprint size) demonstrates a distinct shoulder of elevated polymorphism rate at the aligned edge (**c**), with a significant elevation compared to nearby sequence upstream from the footprint (**d**). DNase I footprints from **a** were aligned to their left edge ($x = 0$) with corresponding polymorphism rates shown (**c**). The increased polymorphism rate cannot be explained by local sequence compositional distortions (**d**). Nucleotide substitution rates in the 11 nucleotides centred on the DNase footprint edge (pink line), and another 11 nucleotides encompassing positions -35 to -25 relative to the footprint edge (green line) were quantified. Darker pink and green filled circles denote the mean of observed substitution rates and lighter

shades denote the mean for the same sites after trinucleotide preserving genomic shuffles. Error bars denote s.d.; statistics by Mann–Whitney test. **e**, Model shows that correlation of increased nucleotide substitution and OJ rates are consistent with increased mutation frequency across heterogeneous DNase I footprints. Polymorphism is reduced at sequence-specific binding sites within the footprints, owing to functional constraint. Therefore, the effect of OF-related mutagenesis in these regions is most sensitively detected in the region immediately adjacent to the binding site (left of vertical dashed blue line, representing footprints aligned to their left edge). This ‘shoulder’ of increased nucleotide substitutions represents sites with increased, OJ-associated mutation is followed by a region of depressed substitution rates, owing to selective effects of the functional binding sites within the footprints (to the right of the dashed blue line). Signals further to the right are not interpretable given the heterogeneity in DNase I footprint sizes. Given strong selection at TF and DNase I footprint sites, this ‘shoulder’ of elevated nucleotide substitutions could represent a measure for the local mutation rate for such regions, analogous to that measured by the fourfold degenerate sites in protein coding sequence.



Extended Data Figure 8 | Model to show Pol- α DNA tract retention downstream of protein binding sites. **a**, OF priming occurs stochastically, with the 5' end of each OF initially synthesized by Pol- α and the remainder of the OF synthesized by Pol- δ . **b**, **c**, OF processing: when Pol- δ encounters the previously synthesized OF, Pol- δ continues to synthesize DNA displacing the 5' end of the downstream OF, which is removed by nucleases to result in mature OFs which are then ligated. The OJs of such mature OFs before ligation were detected previously¹⁷ after depletion of temperature-sensitive

DNA ligase I. They demonstrated that if a protein barrier is encountered (grey circle), Pol- δ progression is impaired, leading to reduced removal of the downstream OF (**b**). Given that $\sim 1.5\%$ of the mature genome is synthesized by Pol- α , a proportion of lagging strands will retain Pol- α -synthesized DNA (red). When Pol- δ progression is impaired by protein binding, this will lead to an increased fraction of fragments containing Pol- α -synthesized DNA downstream of such sites (**c**).

Crystal structure of the V(D)J recombinase RAG1–RAG2

Min-Sung Kim¹, Mikalai Lapkouski^{1†}, Wei Yang¹ & Martin Gellert¹

V(D)J recombination in the vertebrate immune system generates a highly diverse population of immunoglobulins and T-cell receptors by combinatorial joining of segments of coding DNA. The RAG1–RAG2 protein complex initiates this site-specific recombination by cutting DNA at specific sites flanking the coding segments. Here we report the crystal structure of the mouse RAG1–RAG2 complex at 3.2 Å resolution. The 230-kilodalton RAG1–RAG2 heterotetramer is ‘Y-shaped’, with the amino-terminal domains of the two RAG1 chains forming an intertwined stalk. Each RAG1–RAG2 heterodimer composes one arm of the ‘Y’, with the active site in the middle and RAG2 at its tip. The RAG1–RAG2 structure rationalizes more than 60 mutations identified in immunodeficient patients, as well as a large body of genetic and biochemical data. The architectural similarity between RAG1 and the hairpin-forming transposases Hermes and Tn5 suggests the evolutionary conservation of these DNA rearrangements.

To combat the great range of possible infectious agents, the vertebrate immune system deploys a highly diverse population of immunoglobulins and T-cell receptors. In many species this diversity is generated by V(D)J recombination¹. By combinatorial joining of segments of coding sequence, V(D)J recombination is capable of assembling millions of different functional immunoglobulin and T-cell receptor genes^{1,2}. This recombination is initiated by DNA double-strand breaks produced by the RAG1–RAG2 recombinase, at sites flanked by specific recombination signal sequences (RSS). The RSS are of two types, with either 12 or 23 non-conserved nucleotides between conserved heptamer and nonamer modules; one RSS of each type is strictly required for recombination². The two RSS varieties are partitioned so as to focus recombination on V to J, or V to D to J, joining. RAG1 and RAG2 are the only lymphoid-specific factors involved in V(D)J recombination^{3,4}, while the resulting hairpinned coding ends are processed by general repair factors of the non-homologous end-joining pathway^{5,6}.

Since the identification of the *RAG1* and *RAG2* genes^{7,8}, RSS-dependent DNA cleavage by purified RAG1–RAG2 has been reconstituted⁹. RAG1 and RAG2, of 1,040 and 527 residues respectively, cooperate in all their known activities. The catalytic core, regulatory regions, active site residues, DNA-binding domains, two zinc-binding motifs, and some aspects of the interface of RAG1 and RAG2 have been characterized^{3,4}. It was also found that RAG1–RAG2 can function *in vitro* as a transposase^{10,11}, inserting RSS-terminated DNA into a second DNA molecule. Moreover, a large number of human mutations in both RAG proteins that cause severe combined immunodeficiency (SCID) or a milder form known as Omenn syndrome have been identified^{12,13}.

Biochemical and functional studies have shown that portions of RAG1 and RAG2 can be deleted, and the ‘core’ proteins, residues 384–1008 of RAG1 and 1–387 of RAG2, retain targeted cleavage activity *in vitro* and recombination activity (although not fully regulated) in cells^{14–17}. An earlier low-resolution electron microscopic study of the core complex, containing two subunits each of RAG1 and RAG2 bound to a 12RSS and 23RSS DNA pair, revealed the overall shape and localization of RAG proteins¹⁸. Here we report the 3.2 Å crystal structure of the RAG1–RAG2 heterotetramer and its implications for V(D)J recombination.

SEC complex and structure determination

The catalytic cores of mouse RAG1 (384–1008 amino acids) and RAG2 (1–387 amino acids) with maltose binding protein (MBP) fused to their N termini were expressed in HEK293T cells and readily purified (Methods). RAG1–RAG2 was assembled with pre-cleaved 12RSS and 23RSS DNAs in the presence of HMGB1 to form a signal-end complex (SEC)¹⁹, and the purified SEC after removal of the cleaved MBP tags and HMGB1 (Fig. 1a, b) was homogeneous and active in strand transfer (Extended Data Fig. 1).

Crystals were grown over a period of 2–4 weeks (Methods). For phase determination, methionines in the RAG1–RAG2 proteins were substituted by selenomethionine to a level of 40% (Methods). Single-wavelength anomalous diffraction (SAD) data sets of high redundancy were collected at the Se absorption peak from six crystals. Fifty-four of fifty-eight selenium sites were located, together with two Zn²⁺ atoms (one in each RAG1). The electron density map using all SAD data, nominally at 3.7 Å, was superior to that calculated using only the two best sets according to anomalous correlation coefficient²⁰ (Fig. 1c, d and Extended Data Fig. 1a). The heterotetramer of RAG1–RAG2 recombinase was readily traceable (Extended Data Fig. 1b). Although 12RSS and 23RSS DNAs were included in the SEC complex and were also present in dissolved crystals (Extended Data Fig. 1c, d), DNA was not found in the electron density map. Only the four protein chains, with residues 391–1008 of RAG1 and 2–350 of RAG2, were modelled and refined to 3.2 Å (Extended Data Table 1). The carboxy-terminal 37 residues of RAG2 are disordered. In fact, RAG2 (1–351) forms active heterotetramers with RAG1 *in vitro* (Extended Data Fig. 2), and supports V(D)J recombination in cells²¹.

Architecture of RAG1–RAG2

The RAG1–RAG2 crystal structure is remarkably similar to the low-resolution model generated from two-dimensional averaging of negatively stained electron microscopy images¹⁸. It is Y-shaped (125 Å × 150 Å × 90 Å), with the RAG1 dimer forming the bulk and RAG2 situated at the tip of each arm (Fig. 2). There is an oval-shaped gap (40 Å × 60 Å) separating the two RAG1–RAG2 heterodimers (Fig. 2). RAG1 is

¹Laboratory of Molecular Biology, NIDDK, NIH, Bethesda, Maryland 20892, USA. [†]Present address: Department of Cell and Molecular Biology, Karolinska Institute, 171 77 Stockholm, Sweden, and Centre for Structural Systems Biology, DESY, 22607 Hamburg, Germany.

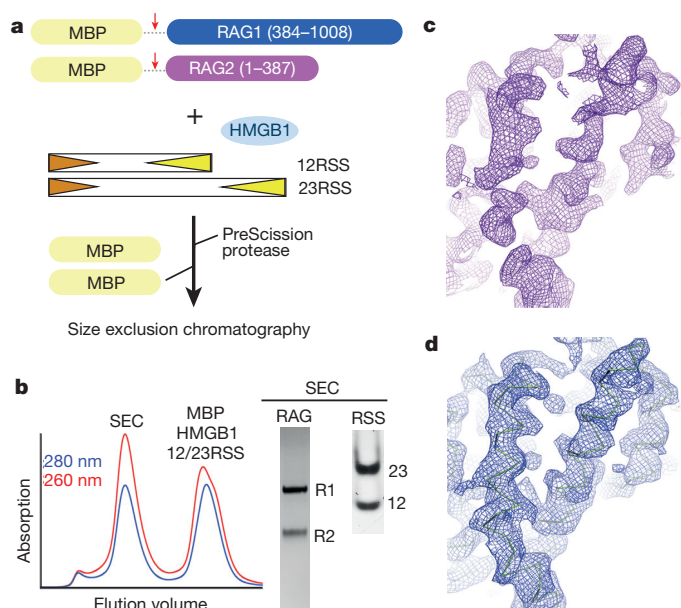


Figure 1 | Structure determination of RAG1–RAG2 recombinase.

a, Procedure of assembling SEC from purified RAG1–RAG2, RSS DNAs and HMGB1. The MBP tags were cleaved off by PreScission protease after SEC formation. **b**, The RAG1–RAG2–DNA complex was purified away from MBP tags, free DNA and HMGB1 on a Superdex-200 column. The eluted peak contains RAG1–RAG2 protein and RSS DNAs, as shown in the protein and DNA denaturing gels stained by Coomassie blue and SYBR green, respectively. **c**, **d**, The experimental electron density map calculated from merging two best SAD data sets (**c**) or all six SAD data sets (**d**).

elongated (100 Å) and composed of seven structural modules (Fig. 3a). The N-terminal nonamer-binding domain (NBD), which superimposes well with the structure determined previously²², forms a domain-swapped and intertwined dimer. The following dimerization and DNA binding domain (DDBD) is connected to the NBD by a flexible linker, and the last helix of the three-helix C-terminal domain (CTD) folds back to complete the DDBD (Fig. 3b), which may be why previous domain dissections failed to isolate this structural entity. Three conserved carboxylates in RAG1 (D600, D708 and E962) have been identified as being essential for catalysis^{23–25}, but it was not clear how the catalytic domain would fold owing to their large separation in the amino-acid sequence. Our structure shows that RAG1 adopts an RNase H fold with an elongated central four-stranded β -sheet, similar to other DDE transposases (named after the three catalytic carboxylates, which are all within the

same RAG1 subunit) (Fig. 3)^{26–28}. Following the DDBD, the extended pre-RNase H (preR) and catalytic RNase H (RNH) domains further lengthen RAG1 to ~100 Å. Each active site is located in the middle of a Y arm, and the two are about 45 Å apart (Fig. 2). Two domains intervene between D708 and E962: ZnC2, which protrudes towards RAG2, and the highly helical ZnH2, which increases the third dimension of RAG1 from 25 Å to 65 Å and eventually brings E962 back to the catalytic centre (Fig. 3a–c). Two regions for Zn^{2+} binding were previously identified^{29,30}. The first (C727, C730) and second (H937 and H942), despite being far apart in the sequence, form one zinc-binding site (Fig. 3c, d) that juxtaposes the catalytic centre, interface with RAG2 (ZnC2), and DNA binding (DDBD and ZnH2) domains.

RAG2 is folded into a six-bladed β -propeller, or Kelch-repeat structure, as predicted^{31,32} (Extended Data Fig. 3). The first N-terminal β -strand belongs to the sixth 4-stranded β -blade and ties the first and last Kelch repeats together. Compared with other β -propeller structures, the six blades of the doughnut-shaped RAG2 are more distorted in planarity and spacing between blades (Extended Data Fig. 3). These distortions are most notable at the interface with RAG1 (Fig. 4a). As usual for β -propeller proteins, one face of RAG2 has extended loops. Many loops, particularly those connecting adjacent blades or the middle two strands of each blade, are involved in interacting with the preR, RNH and ZnC2 domains of RAG1 (Fig. 4b, c). The interface between RAG1 and RAG2 is highly conserved from fish to humans, encompassing both polar and hydrophobic interactions and dovetailed with ridges and canyons (Fig. 4). Interestingly, RAG2 contacts RAG1 near the active site³³, including E607 and V615 (connected by a disordered loop) and E719 to V724 contacted by the long RAG2 loop 335–339 and residue R39, respectively (Figs 2 and 4c). It is likely that in the presence of DNA substrate, RAG2 assists in formation of the catalytic site and DNA cleavage.

SCID and Omenn syndrome mutations

Over 60 missense mutations leading to SCID or Omenn syndrome, a milder type of immunodeficiency, due to defective DNA processing in V(D)J recombination have been mapped to the catalytic cores of RAG1 and RAG2 (refs 12, 13). The disease-involved residues are identical between human and mouse RAG proteins, except for M435 in human RAG1 being replaced by L432 in mouse. Residue numbers in mouse and human RAG1 differ by three (Extended Data Table 2). For consistency, all residues are numbered here according to the mouse protein that we are studying. The SCID and Omenn syndrome mutations fall into four classes. The first class of mutations clearly destabilizes the tertiary structure of RAG1–RAG2. For example, mutations of the zinc-binding site, C727E and the adjacent L729F (Fig. 3d), would perturb the structure

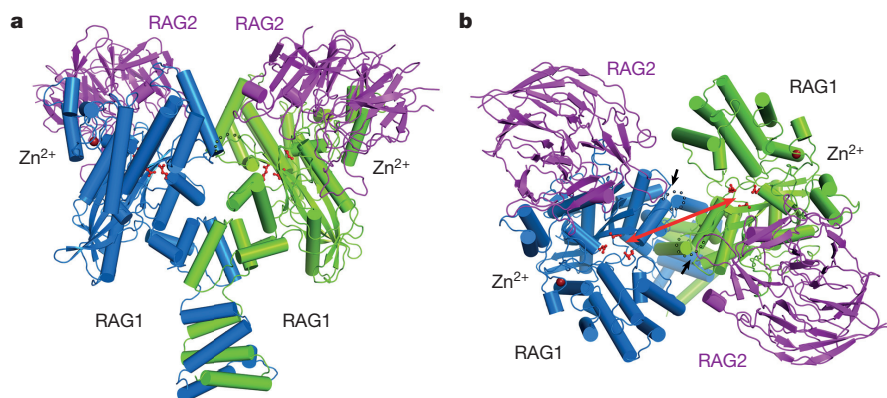


Figure 2 | Crystal structure of RAG1–RAG2. **a**, **b**, Front (**a**) and top (**b**) view of the RAG1–RAG2 heterotetramer. The two RAG1 chains are shown in blue and green ribbon diagrams, and both RAG2 subunits are shown in magenta. The active sites are highlighted by the three carboxylates shown as red

sticks. The zinc ions are shown as dark red spheres. The distance between the two active sites is ~45 Å (marked by the red double arrowheads). The disordered loop of residues 608–614 in RAG1 near the RAG1–RAG2 interface is marked by a dotted line and grey arrowhead.

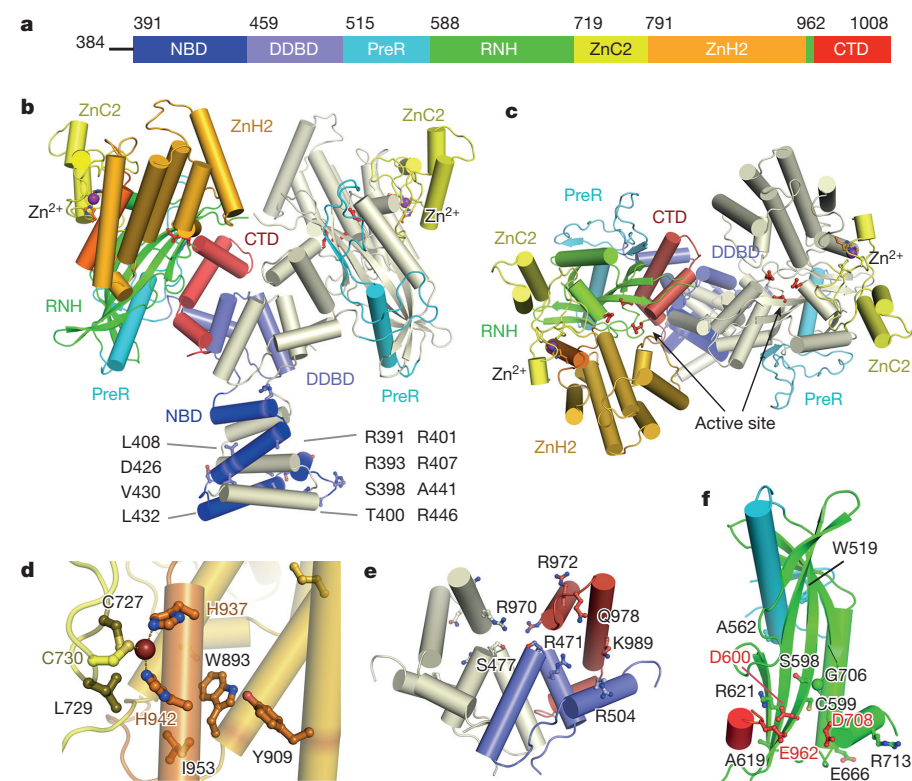


Figure 3 | The RAG1 structure. **a**, A diagram of RAG1 domains with boundaries indicated by residue numbers. **b**, Cartoons of RAG1 dimer. One subunit is colour-coded as in **a**, and the other in silver except for the preR and ZnC2 domains. **c**, An orthogonal view of **b**. **d**, The Zn^{2+} coordination by two Cys (of ZnC2) and two His residues (of ZnH2). **e**, The DDBD and CTD domains. **f**, The RNH and a portion of preR domain with the catalytic carboxylates shown in red sticks. In panels **b**, **d**–**f**, SCID/Omenn syndrome mutations are shown as coloured sticks with black labels.

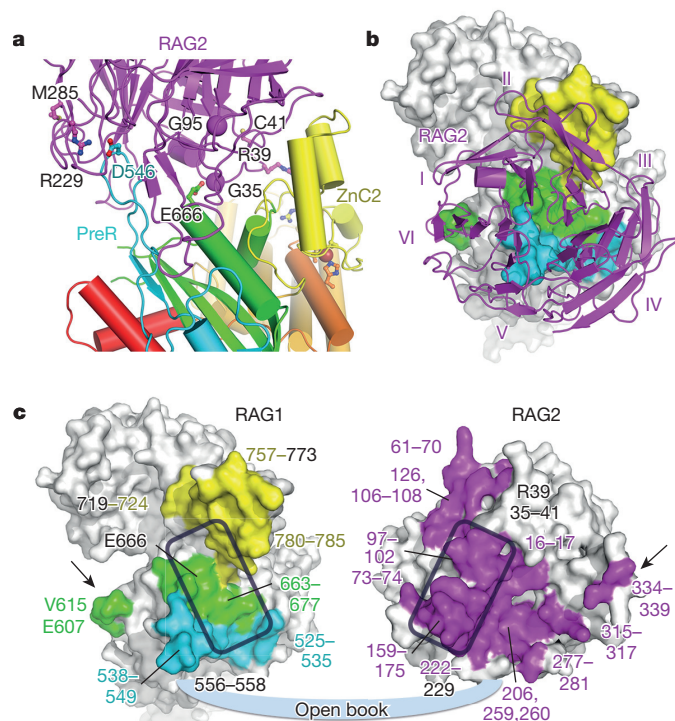


Figure 4 | The interface between RAG1 and RAG2. **a**, One side of the doughnut-shaped RAG2 interacts with the preR, RNH and ZnC2 domains of RAG1 (colour-coded as in Fig. 3). The SCID/Omenn syndrome mutations in RAG2 are shown and labelled in black. **b**, An orthogonal view of the RAG1–RAG2 interface. The Kelch repeats of RAG2 are labelled I to VI. **c**, Based on **b**, the interface of RAG1 and RAG2 is shown as an open book. The regions at the interface are indicated according to the colour code, and SCID/Omenn syndrome mutations are labelled in black. The mirrored arrowheads and boxes indicate matching surfaces.

of the ZnC2, ZnH2 and adjoining domains. Three immunodeficiency mutations—W893R, Y909C and I953R—probably destabilize the hydrophobic core next to the zinc-binding site (Fig. 3d), reflecting the importance of this region in the overall structure of RAG1–RAG2. A number of laboratory-generated mutations with loss of function also are likely to destabilize the structure of certain domains in RAG1, for example W893A.

The second class includes polar residues exposed to solvent and concentrated in two areas likely for DNA binding. Seven out of nine immunodeficiency mutations found in the DDBD and CTD domains are at a cluster of polar residues (Arg, Lys, Ser or Gln) (Fig. 3e). Systematic mutation of positively charged residues in RAG1 (ref. 34) also showed that K966, R969, R977 and H990 in CTD contribute to DNA binding and cleavage. In parallel, 12 SCID/Omenn syndrome mutations in the NBD domain (Fig. 3b, Extended Data Fig. 4), which overlap with many highly conserved non-polar and polar residues, are involved in structural integrity and sequence-specific binding to the nonamer DNA²².

The third class of mutations is clustered around the active site (Fig. 3f). Some may alter the structure of the catalytic centre (S598P, C599W, A619P, R696Q/W, G706D), and others may change its DNA-binding properties (E666G, R621C/H, R713W). It is not surprising that engineered mutations of conserved polar residues surrounding the active site (E597, E709, D792, H795 and E959) lead to defects in DNA cleavage^{25,34}.

The last class of SCID/Omenn syndrome mutations is located at the interface of RAG1 and RAG2. Four of six disease mutations in the RAG2 core are concentrated at the subunit interface (Fig. 4). G35V, R39G and C41W are at the interface with the ZnC2 and RNH domains of RAG1. R229 forms salt bridges with D546 (preR) of RAG1, which has been accurately mapped to the interface with RAG2 (ref. 35). G95R is at the base of a long loop that reaches to the RNH domain and sandwiches E666 (which itself is mutated in people with SCID) with G35 (Fig. 4a). The importance of the subunit interface is also evident in the disease mutations located on the RAG1 side (R556S, R558C/H, E666G and R773Q) and additional mutations identified in laboratories (E719Q and R773A)^{25,34} (Extended Data Table 3).

A model of the RAG1–RAG2–RSS DNA complex

Structures of five DDE transposases complexed with DNA have been reported^{36–40}. They can be segregated into two groups based on the markedly different DNA orientations relative to the catalytic dimer (Extended Data Figs 5 and 6). Hermes (a member of the hAT transposase family) and Tn5, both of which generate a hairpin intermediate during DNA processing, belong to the same group, and both contain an α -helix extending from one catalytic centre (by contributing the last catalytic residue Glu) to contact the DNA bound to the other subunit (Extended Data Fig. 5). Very similar architectural features are found in the RAG1 dimer. After superimposition of the RNH domain of RAG1 and Hermes (Fig. 5a), not only does the 16-base-pair (bp) DNA co-crystallized with Hermes fit into the active site of RAG1, but the second DNA of Hermes is close to the other RAG1 catalytic centre (Fig. 5b). The first 7 bp of the DNA may mimic the heptamer of each RSS (Fig. 5b). Remarkably, the two DNAs modelled in RAG1 are connected by the α -helix (964–975 amino acids) that immediately follows the catalytic residue E962, just as is found in Hermes and Tn5. None of the aromatic residues previously suggested as functioning in DNA hairpin formation^{41,42} is situated near the active site. G851 and N852 in the ZnH2 domain of RAG1 appear to replace W319 in the equivalent helical insertion domain of Hermes that stacks on the 3'-end DNA base (Fig. 5b). In corroboration of this model, mutations of N852 or residues 970–978 in the CTD are implicated in Omenn syndrome (Figs 3e and 5b).

The rest of the RSS DNA can also be modelled based on the published crystal structure of an isolated NBD bound to a 12-bp nonamer DNA²² (Extended Data Fig. 4). From DDBD to CTD, the two halves of the RAG1–RAG2 tetramer are rather symmetric (Fig. 5c), but the NBD domains are not related by the same dyad. Although perfectly symmetrical internally, the intertwined NBDs are tilted relative to the rest of the protein. As a result, the two nonamer DNAs bound to the NBDs

would be oriented differently to the two catalytic centres. One is very close to the 16-bp DNA modelled into the RNase H domain (Fig. 5c). Notably, the sum of the two DNA segments is about 28 bp, the total length of 12RSS DNA. In contrast, the other pair of DNA segments is separated by ~ 30 Å, which may mimic the 23RSS with its additional 11 bp connecting the nonamer and heptamer (see Supplementary Video 1). In this model, the nonamer and heptamer ends of each RSS interact with different RAG1 subunits in a *trans* configuration, as mutational study has suggested²⁷. A sharp kink is unavoidable in each RSS DNA as modelled (Fig. 5c), and HMGB1 could stabilize such kinks to facilitate the gene rearrangement^{43,44}.

The surface of RAG1–RAG2 traversed by the modelled DNAs is both highly positively charged and highly conserved (Extended Data Fig. 7). The only exception is the NBD, which is not as highly charged as the DDBD and ZnH2 regions. This may correlate with the sequence-specific recognition of the nonamer. Beyond the regions that are modelled to bind RSS DNAs, extensive surface areas along the rim of the Y arms from RAG2 to ZnH2 are positively charged and partially conserved (Supplementary Video 1). These areas could bind up to 20 bp of coding DNA flanking the RSS. Although 6-bp coding flanks can be slowly cleaved by RAG1–RAG2, efficient cleavage requires more than 15 bp of flanking DNA⁴⁵. Interactions of the coding flanks with the top of the RAG1–RAG2 complex may explain why many mutations in the RAG1 and RAG2 interface have an impact on DNA cleavage³³.

Concluding remarks

The structure of RAG1–RAG2 reveals the architecture of the complex and the composition of its functional sites. It rationalizes the effects of many mutations associated with human immune deficiencies. Evolutionarily, eukaryotic hAT transposases and RAG1–RAG2 recombinase, which cleave DNA duplex in two steps and leave a hairpin on the flanking DNA, are thought to be rather different from the bacterial transposases that cleave DNA in three steps and form a hairpin intermediate on the recognition DNA (Fig. 5d). The similar enzyme–substrate association found in Hermes and Tn5 and their structural relationship with RAG1 has led us to propose that the two DNA recombination processes are identical in mechanism and configuration, differing only in the nucleophiles used at each step, a water molecule versus a 3'-OH.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 October; accepted 22 December 2014.

Published online 18 February 2015.

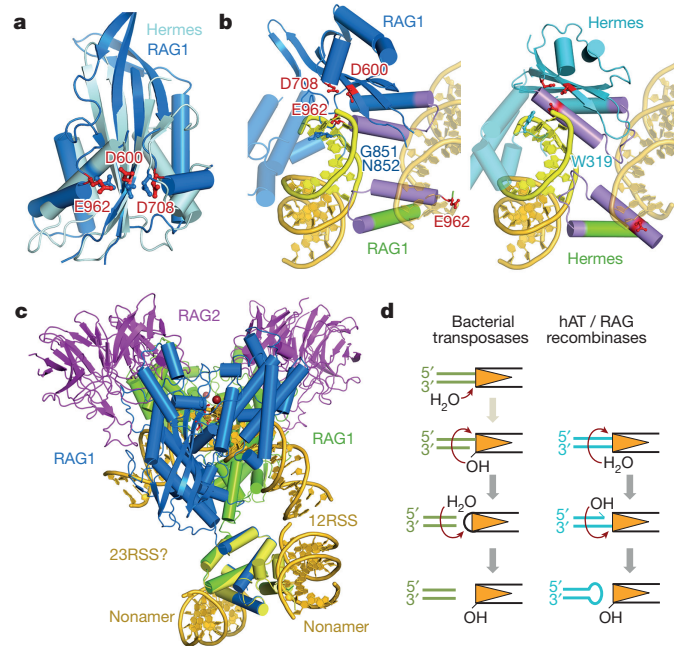


Figure 5 | A RAG1–RAG2–DNA model. **a**, Superposition of the RNH domains of RAG1 (red active site) and Hermes (Protein Data Bank (PDB) 4D1Q (ref. 37)). **b**, The superposition places the 16-bp DNA of the Hermes–DNA complex in the RAG1 active site. The DNA is coloured yellow (the first 7 bp) and gold. The second DNA is shown without additional manipulation. The α -helices bridging the two DNAs are coloured lilac. **c**, The RAG1–RAG2–RSS DNA model resulting from superposition with the Hermes and the NBD–DNA complex (12 bp, PDB 3GNA)²². **d**, DNA cleavage by hairpin-forming bacterial and eukaryotic transposases. The recognition sequences are represented by orange triangles.

1. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* **280**, 288–294 (1979).
2. Lewis, S. M. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol.* **56**, 27–150 (1994).
3. Gellert, M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu. Rev. Biochem.* **71**, 101–132 (2002).
4. Schatz, D. G. & Swanson, P. C. V(D)J recombination: mechanisms of initiation. *Annu. Rev. Genet.* **45**, 167–202 (2011).
5. Deriano, L. & Roth, D. B. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* **47**, 433–455 (2013).
6. Boboila, C., Alt, F. W. & Schwer, B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv. Immunol.* **116**, 1–49 (2012).
7. Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* **248**, 1517–1523 (1990).
8. Schatz, D. G., Oettinger, M. A. & Baltimore, D. The V(D)J recombination activating gene, RAG-1. *Cell* **59**, 1035–1048 (1989).
9. McBlane, J. F. *et al.* Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* **83**, 387–395 (1995).
10. Agrawal, A., Eastman, Q. M. & Schatz, D. G. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394**, 744–751 (1998).
11. Hiom, K., Melek, M. & Gellert, M. DNA transposition by the RAG1 and RAG2 proteins: A possible source of oncogenic translocations. *Cell* **94**, 463–470 (1998).

12. Lee, Y. N. *et al.* A systematic analysis of recombination activity and genotype-phenotype correlation in human recombination-activating gene 1 deficiency. *J. Allergy Clin. Immunol.* **133**, 1099–1108 (2014).
13. Piirilä, H., Valiaho, J. & Vihinen, M. Immunodeficiency mutation databases (IDbases). *Hum. Mutat.* **27**, 1200–1208 (2006).
14. Cuomo, C. A. & Oettinger, M. A. Analysis of regions of RAG-2 important for V(D)J recombination. *Nucleic Acids Res.* **22**, 1810–1814 (1994).
15. Sadofsky, M. J., Hesse, J. E. & Gellert, M. Definition of a core region of RAG-2 that is functional in V(D)J recombination. *Nucleic Acids Res.* **22**, 1805–1809 (1994).
16. Sadofsky, M. J., Hesse, J. E., McBlane, J. F. & Gellert, M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res.* **21**, 5644–5650 (1993).
17. Silver, D. P., Spanopoulou, E., Mulligan, R. C. & Baltimore, D. Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. *Proc. Natl Acad. Sci. USA* **90**, 6100–6104 (1993).
18. Grundy, G. J. *et al.* Initial stages of V(D)J recombination: the organization of RAG1/2 and RSS DNA in the postcleavage complex. *Mol. Cell* **35**, 217–227 (2009).
19. Jones, J. M. & Gellert, M. Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends. *Proc. Natl Acad. Sci. USA* **98**, 12926–12931 (2001).
20. Liu, Q., Zhang, Z. & Hendrickson, W. A. Multi-crystal anomalous diffraction for low-resolution macromolecular phasing. *Acta Crystallogr. D* **67**, 45–59 (2011).
21. Gigi, V. *et al.* RAG2 mutants alter DSB repair pathway choice in vivo and illuminate the nature of 'alternative NHEJ'. *Nucleic Acids Res.* **42**, 6352–6364 (2014).
22. Yin, F. F. *et al.* Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. *Nature Struct. Mol. Biol.* **16**, 499–508 (2009).
23. Fugmann, S. D., Villey, I. J., Ptaszek, L. M. & Schatz, D. G. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. *Mol. Cell* **5**, 97–107 (2000).
24. Kim, D. R., Dai, Y., Mundy, C. L., Yang, W. & Oettinger, M. A. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes Dev.* **13**, 3070–3080 (1999).
25. Landree, M. A., Wibbenmeyer, J. A. & Roth, D. B. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes Dev.* **13**, 3059–3069 (1999).
26. Nesmelova, I. V. & Hackett, P. B. DDE transposases: Structural similarity and diversity. *Adv. Drug Deliv. Rev.* **62**, 1187–1195 (2010).
27. Swanson, P. C. The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. *Mol. Cell. Biol.* **21**, 449–458 (2001).
28. Yang, W. & Steitz, T. A. Recombining the structures of HIV integrase, RuvC and RNase H. *Structure* **3**, 131–134 (1995).
29. Gwyn, L. M., Peak, M. M., De, P., Rahman, N. S. & Rodgers, K. K. A zinc site in the C-terminal domain of RAG1 is essential for DNA cleavage activity. *J. Mol. Biol.* **390**, 863–878 (2009).
30. Rodgers, K. K. *et al.* A zinc-binding domain involved in the dimerization of RAG1. *J. Mol. Biol.* **260**, 70–84 (1996).
31. Aravind, L. & Koonin, E. V. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *J. Mol. Biol.* **287**, 1023–1040 (1999).
32. Callebaut, I. & Mornon, J. P. The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis. *Cell. Mol. Life Sci.* **54**, 880–891 (1998).
33. Swanson, P. C. & Desiderio, S. RAG-2 promotes heptamer occupancy by RAG-1 in the assembly of a V(D)J initiation complex. *Mol. Cell. Biol.* **19**, 3674–3683 (1999).
34. Huye, L. E., Purugganan, M. M., Jiang, M. M. & Roth, D. B. Mutational analysis of all conserved basic amino acids in RAG-1 reveals catalytic, step arrest, and joining-deficient mutants in the V(D)J recombinase. *Mol. Cell. Biol.* **22**, 3460–3473 (2002).
35. Ko, J. E., Kim, C. W. & Kim, D. R. Amino acid residues in RAG1 responsible for the interaction with RAG2 during the V(D)J recombination process. *J. Biol. Chem.* **279**, 7715–7720 (2004).
36. Hare, S., Gupta, S. S., Valkov, E., Engelman, A. & Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**, 232–236 (2010).
37. Hickman, A. B. *et al.* Structural basis of hAT transposon end recognition by Hermes, an octameric DNA transposase from *Musca domestica*. *Cell* **158**, 353–367 (2014).
38. Montañó, S. P., Pigli, Y. Z. & Rice, P. A. The mu transpososome structure sheds light on DDE recombinase evolution. *Nature* **491**, 413–417 (2012).
39. Richardson, J. M., Colloms, S. D., Finnegan, D. J. & Walkinshaw, M. D. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell* **138**, 1096–1108 (2009).
40. Steiniger-White, M., Rayment, I. & Reznikoff, W. S. Structure/function insights into Tn5 transposition. *Curr. Opin. Struct. Biol.* **14**, 50–57 (2004).
41. Grundy, G. J., Hesse, J. E. & Gellert, M. Requirements for DNA hairpin formation by RAG1/2. *Proc. Natl Acad. Sci. USA* **104**, 3078–3083 (2007).
42. Lu, C. P., Sandoval, H., Brandt, V. L., Rice, P. A. & Roth, D. B. Amino acid residues in RAG1 crucial for DNA hairpin formation. *Nature Struct. Mol. Biol.* **13**, 1010–1015 (2006).
43. Aidinis, V. *et al.* The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1–RAG2. *Mol. Cell. Biol.* **19**, 6532–6542 (1999).
44. van Gent, D. C., Hiom, K., Paull, T. T. & Gellert, M. Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO J.* **16**, 2665–2670 (1997).
45. Kim, D. R. & Oettinger, M. A. Functional analysis of coordinated cleavage in V(D)J recombination. *Mol. Cell. Biol.* **18**, 4679–4688 (1998).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Grundy and S. Ramon-Maiques for the pioneering work that made this study possible, and D. Leahy for editing the manuscript. The research was supported by the intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

Author Contributions M.-S.K. prepared SEC complexes and carried out crystallography. M.L. developed protocols for protein expression and initial crystallization. W.Y. and M.G. designed the project, and M.-S.K., W.Y. and M.G. prepared the manuscript.

Author Information Atomic coordinates and structure factors have been deposited with the Protein Data Bank under accession code 4WWX. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.Y. (wei.yang@nih.gov) or M.G. (gellert@helix.nih.gov).

METHODS

No statistical methods were used to predetermine sample size.

Protein expression and purification. Mouse core RAG1 (384–1008 amino acids) and RAG2 (1–387 amino acids) were cloned into the pLEXm-based⁴⁶ mammalian expression vector, modified with an N-terminal His₆ tag followed by MBP tag and a PreScission cleavage site (LEVLFQ/GP) (where ‘/’ indicates the cleavage site). The N-terminal Met of RAG2 was mutated to Val (M1V) during cloning. To express the RAG1–RAG2 complex in HEK293T cells, 500 µg of each of the RAG1 and RAG2 expression plasmids were mixed with 4 mg of polyethylenimine (Polysciences) in 35 ml of Hybridoma medium (Invitrogen) to transfect 1 l of HEK293T cells grown in suspension culture in Freestyle 293 medium (Invitrogen), supplemented with 1% FBS when the cell density reached 1.5 million per millilitre. Four days after transfection, cells were harvested and stored at –80 °C. Cell paste (~8 g from 1 l culture) was re-suspended in 50 ml of lysis buffer containing 20 mM HEPES (pH 7.3), 1 M KCl, 1 mM Tris (2-carboxyethyl) phosphine (TCEP) (pH 7.0), 1 mM EDTA and protease inhibitor cocktail (Roche), and lysed by sonication. After centrifuging at 35,000 r.p.m. for 1 h, the clarified lysate was mixed with 5 ml of amylose resin (NEB), which was pre-equilibrated with the lysis buffer and incubated with rotation for 1 h. After pouring the resin into a column and washing thoroughly with 200 resin volumes of the lysis buffer, the RAG1–RAG2 protein was eluted by an amylose elution buffer containing 20 mM HEPES (pH 7.3), 500 mM KCl, 40 mM maltose and 1 mM TCEP. The eluted protein, which consisted mainly of RAG1–RAG2 heterotetramer⁴⁷ was concentrated and stored at –80 °C after adding glycerol to 20% final concentration. In contrast to RAG1 and RAG2 core proteins expressed in insect cells, which required activity-based purification for structural studies¹⁸, RAG1–RAG2 expressed in human cells is highly active. Human HMGB1 (1–163 amino acids) was prepared as reported previously^{18,19}.

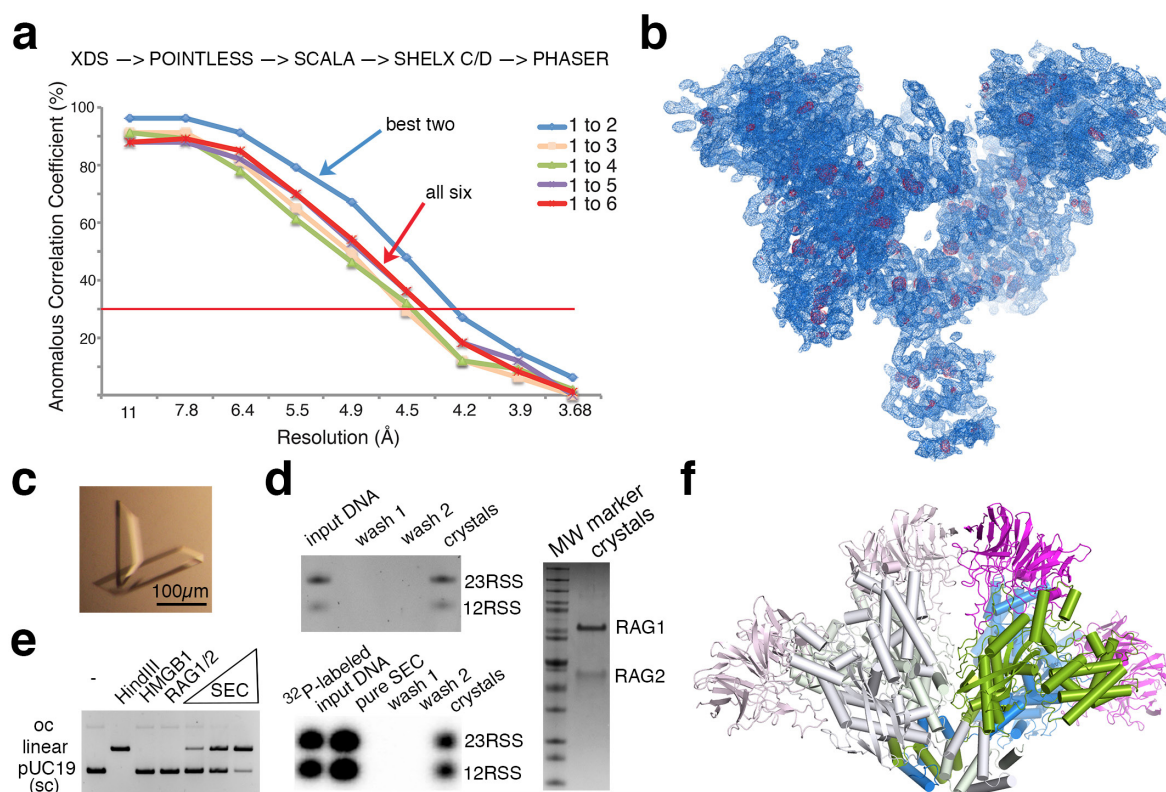
To make the SEC complex, purified RAG1, RAG2, 12RSS and 23RSS DNAs were mixed at 2:2:1:1 molar ratio in the presence of a 2-fold excess of HMGB1 with the buffer of 20 mM HEPES (pH 7.3), 150 mM KCl, 5 mM CaCl₂, and 1 mM TCEP and incubated for 1 h at 37 °C. After removing the MBP tag by addition of PreScission protease (1:100 mass ratio of protease to substrate) overnight at 4 °C, further purification was performed using gel filtration (Superdex 200, GE Healthcare) in 20 mM HEPES (pH 7.3), 500 mM KCl, 5 mM CaCl₂ and 1 mM TCEP, which removed HMGB1 along with free DNA, the MBP tag, and PreScission protease. All purification steps were performed at 4 °C. HMGB1 could be retained with RAG1–RAG2 and RSS DNA if KCl concentration was reduced to below 100 mM (Extended Data Fig. 2), but we were unable to crystallize SEC completed with HMGB1. To prepare selenomethionine (SeMet)-labelled RAG1–RAG2 complex, HEK293T cells were transfected after transfection to methionine-free Freestyle 293 medium (Invitrogen) supplemented with 25 mg l⁻¹ L-SeMet (Acros Organics) and 1% dialysed FBS (Invitrogen). Three days later, cells were collected and protein was purified in the same way as native protein. Mass spectrometry analysis of trypsin-digested SeMet-labelled RAG1–RAG2 peptides was performed at the Taplin Mass Spectrometry Facility (taplin.med.harvard.edu). It showed that about 40% of methionines were substituted by SeMet.

Crystallization and data collection. Crystals of the RAG1–RAG2–DNA complexes were grown by the hanging-drop vapour diffusion method at 4 °C over 3 weeks. Equal volumes of protein (~5 mg ml⁻¹) and reservoir solution containing 100 mM MES (pH 7.1), 10–15% PEG 3350, 200 mM tribasic ammonium citrate (pH 7.0) and 100 mM KCl were mixed in each droplet. Crystals were cryo-protected in reservoir solution supplemented with 25% ethylene glycol and flash frozen in liquid nitrogen. We were able to crystallize tetrameric RAG1–RAG2 alone as well as RAG1–RAG2 with a single 12RSS or 23RSS, but these crystals were small, and none of them diffracted X-rays as well as the SEC complex. Crystals of SeMet-labelled RAG1–RAG2 complex were grown under similar conditions. Native and SeMet-labelled complex both crystallized in the C222₁ space group with two RAG1 and two RAG2 (one RAG1/2 heterotetramer) in each asymmetric unit. Data were collected at 100 K for native and SeMet-derivative crystals at beam lines 22ID and 23ID of the Advanced

Photon Source (APS) at Argonne National Laboratory. All data were indexed, integrated and scaled with the XDS package⁴⁸ (Extended Data Table 1).

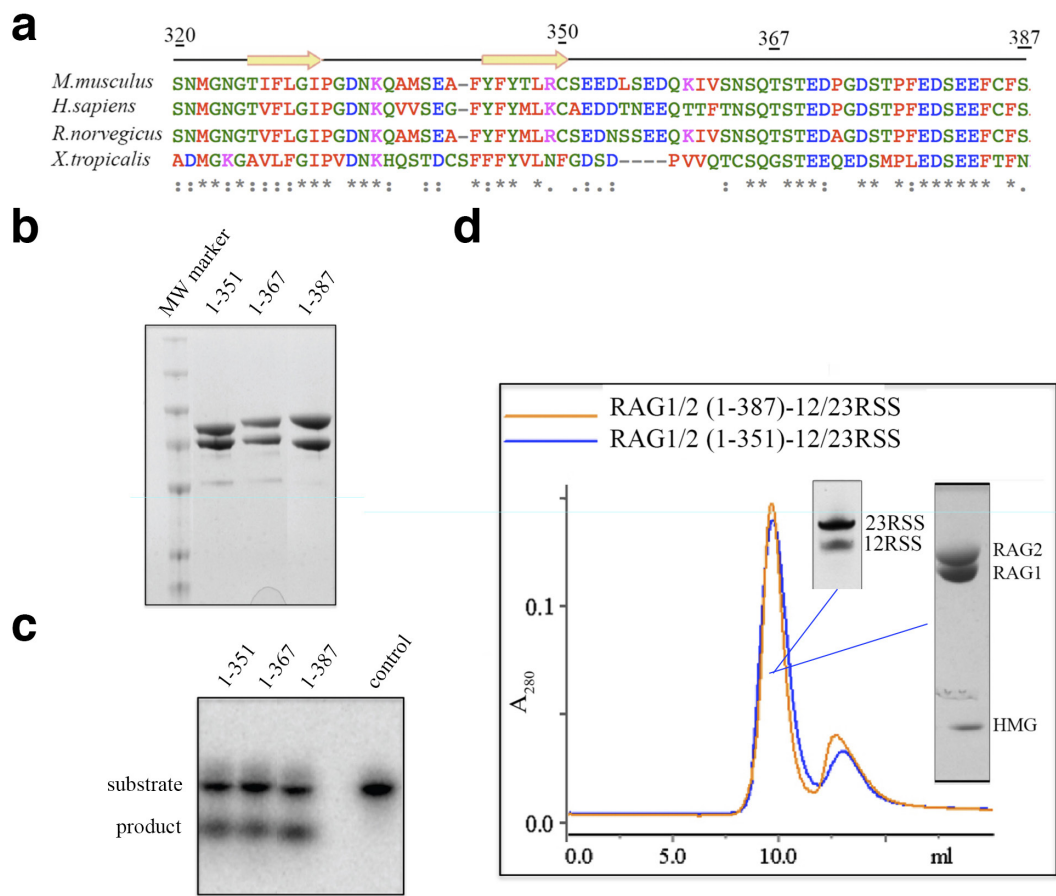
Structure determination and refinement. Phases were determined by the single-wavelength anomalous diffraction (SAD) method and multi-crystal averaging²⁰. SAD data were collected from six SeMet-substituted crystals with the best resolution of 3.7 Å (Extended Data Table 1). Data were processed according to the published procedure²⁰ (Extended Data Fig. 1a). Selenium sites were identified using SHELXD⁴⁹ and refined with PHASER⁵⁰. Out of the 58 highest anomalous peaks, 54 corresponded to selenium sites, and two were Zn²⁺ ions. Phases were improved by density modification using RESOLVE⁵¹ and the overall figure-of-merit was 0.79. The RAG1–RAG2 model was built manually in COOT⁵². Although the experimental electron density map contained breaks in the main chains, and side-chain definition was not perfect, the register of the polypeptide chains was readily determined based on the SeMet sites. This initial model was refined in Phenix⁵³ and manually improved using COOT. Secondary structure restraints and non-crystallographic twofold symmetry averaging restraints were used throughout the refinement. The RAG1–RAG2 structure was refined to 3.2 Å with R_{work} and R_{free} of 20.6% and 25.9%, respectively (Extended Data Table 1). The quality of the structure was validated with MolProbity⁵⁴. 90.7% of residues are in the favoured regions of the Ramachandran plot, 9.3% in additional allowed regions, and no residue in the disallowed region. The final model contains amino acids 391–1008 of RAG1 and amino acids 2–350 of RAG2, and one Zn²⁺ ion in each RAG1. The N-terminal residues (391–404) are ordered in one RAG1 subunit. Owing to poor electron densities, residues 608–616 of RAG1 and 82–88, 242–244, 254–255 and 334–337 of RAG2 were not included in the final model. Crystal packing of two neighbouring RAG1–RAG2 tetramers appears to occlude one nonamer-binding site in each RAG1–RAG2 complex (Extended Data Fig. 1e). No water molecules were added. All structure figures were prepared with PyMOL (<http://www.pymol.org>), and sequence conservation analysis was performed using ClustalW⁵⁵.

46. Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D* **62**, 1243–1250 (2006).
47. Swanson, P. C. A RAG-1/RAG-2 tetramer supports 12/23-regulated synapsis, cleavage, and transposition of V(D)J recombination signals. *Mol. Cell. Biol.* **22**, 7790–7801 (2002).
48. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
49. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122 (2008).
50. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
51. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
52. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
53. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
54. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
55. Chenna, R. *et al.* Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497–3500 (2003).
56. Schumacher, F. R., Sorrell, F. J., Alessi, D. R., Bullock, A. N. & Kurz, T. Structural and biochemical characterization of the KLHL3-WNK kinase interaction important in blood pressure regulation. *Biochem. J.* **460**, 237–246 (2014).
57. Corbett, K. D., Shultzaberger, R. K. & Berger, J. M. The C-terminal domain of DNA gyrase A adopts a DNA-bending β-pinwheel fold. *Proc. Natl Acad. Sci. USA* **101**, 7293–7298 (2004).
58. Schuetz, C. *et al.* Lesson from hypomorphic recombination-activating gene (RAG) mutations: Why asymptomatic siblings should also be tested. *J. Allergy Clin. Immunol.* **133**, 1211–1215 (2014).
59. Corneo, B. *et al.* Identical mutations in RAG1 or RAG2 genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome. *Blood* **97**, 2772–2776 (2001).
60. Dhingra, N. *et al.* Severe combined immunodeficiency caused by a new homozygous RAG1 mutation with progressive encephalopathy. *Hematol. Oncol. Stem Cell Ther.* **7**, 44–49 (2014).



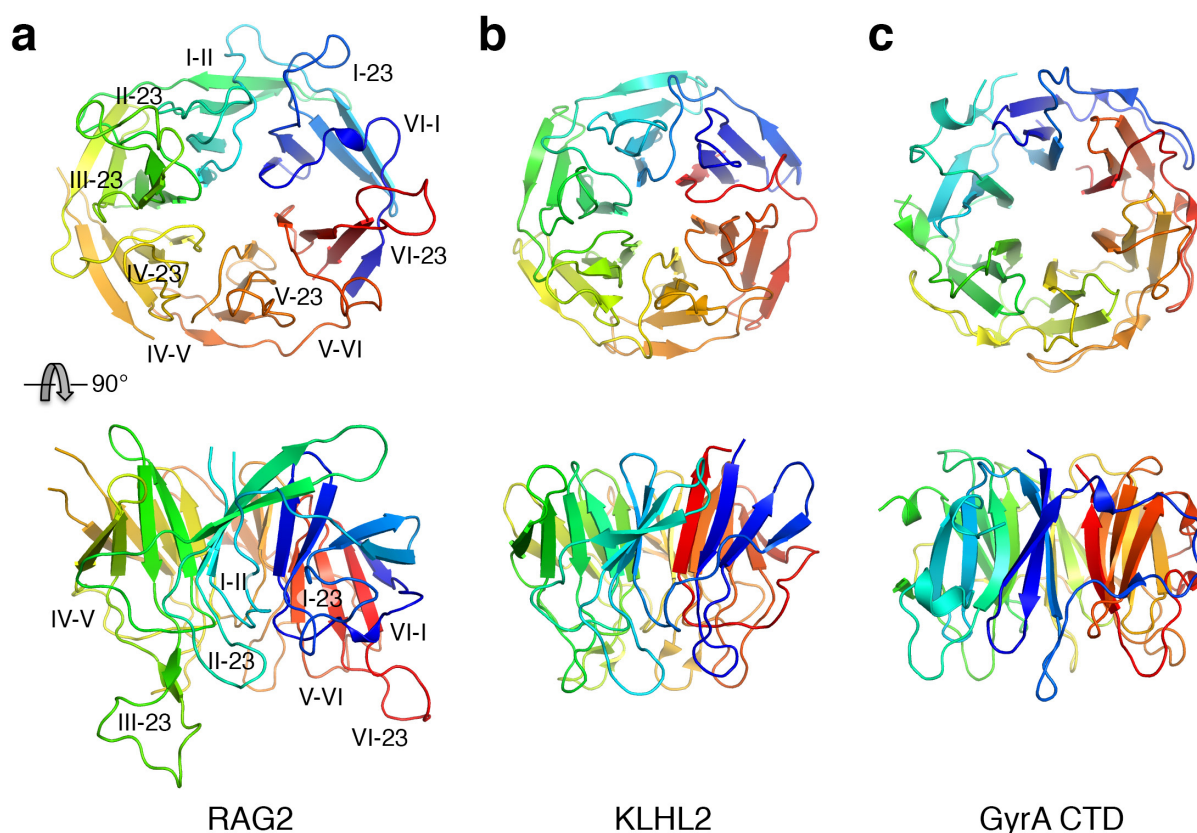
Extended Data Figure 1 | Structure determination. **a**, The plot of correlation coefficient (CC) of anomalous signal versus resolution. The red line indicates the cutoff of CC = 0.3. Merging data from the two best crystals produced a better CC than merging data from all six crystals. The data processing procedure is outlined above the plot²⁰. **b**, The SAD experimental map contoured at 1.3σ showed the content of an asymmetric unit. The Se anomalous map is contoured at 3.0σ in red. **c**, A typical crystal of RAG1–RAG2. **d**, The content of crystals was examined by protein and DNA denaturing gels after a thorough wash of the crystals and stained by Coomassie blue and SYBR green. To confirm the 1:1 molar ratio of 12 and 23RSS DNA,

³²P-labelled input RSS DNAs and those in SEC complexes before and after crystallization are shown beneath the SYBR-green-stained DNA gel. **e**, Transposition assay of the purified SEC (RAG1–RAG2–12/23RSS DNA complex) used for crystallization. Supercoiled pUC19 (sc, with a small amount of open circle, oc) was the target; it was linearized by HindIII as a control. The SEC (0.25, 0.5 and 1.0 μM) was active in concerted transposition and thus linearizing pUC19. In contrast, RAG1–RAG2 or HMGB1 (0.5 μM) each alone was not active. **f**, Crystal packing of neighbouring RAG1–RAG2 complexes (shown in dark and light colours) occludes one nonamer-binding site in each heterotetramer of RAG1–RAG2.



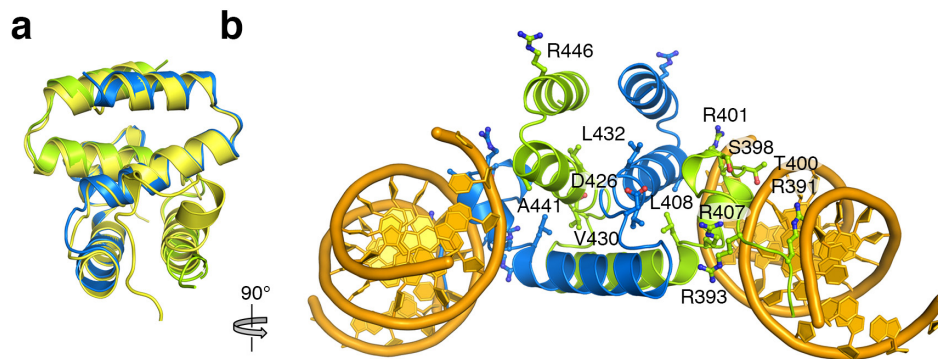
Extended Data Figure 2 | RAG2 core fragment (1–351 amino acids) is active. **a**, Sequence alignment of RAG2 from mouse (320–387 amino acids), human, rat and *Xenopus* with predicted secondary structures shown above. **b**, Core RAG2 (1–387) and two further truncated RAG2 variants (1–351 and 1–367) were constructed with a non-cleavable N-terminal MBP tag and co-expressed with the tag-less core RAG1. The Coomassie blue R-250 stained SDS gel shows the purified RAG1–RAG2 complexes. **c**, Purified RAG1–RAG2 complexes with truncated RAG2 variants are equally active in cleaving a 32 P-labelled 12RSS DNA (in the presence of a 23RSS and Mg^{2+} , as examined

by TBE-Urea gel). **d**, Elution profiles of RAG1–RAG2 (both long and short forms) complexed with DNA from Superdex-200 (S200) in a low salt buffer (50 mM HEPES pH 7.0, 60 mM KCl, 1 mM maltose and 2 mM DTT). Regardless of the length of RAG2, the major S200 eluant peak came out at the same time point and contained RAG1, RAG2 (1–351 or 1–387) and HMGB1 proteins, as shown in the SDS gel (right insert), as well as 12 and 23RSS oligonucleotides, as confirmed by a TBE-Urea gel stained by SYBR green (left insert).



Extended Data Figure 3 | Comparison of RAG2 with β -propeller and β -pinwheel structures. KLHL2 (PDB 4CHB)⁵⁶ is selected to represent the β -propeller proteins, and the C-terminal domain (CTD) of GyrA (PDB 1SUU)⁵⁷ is selected to represent the β -pinwheel structures. After superposition, RAG2 (a), KLHL2 (b) and GyrA (c) are shown side-by-side individually in two

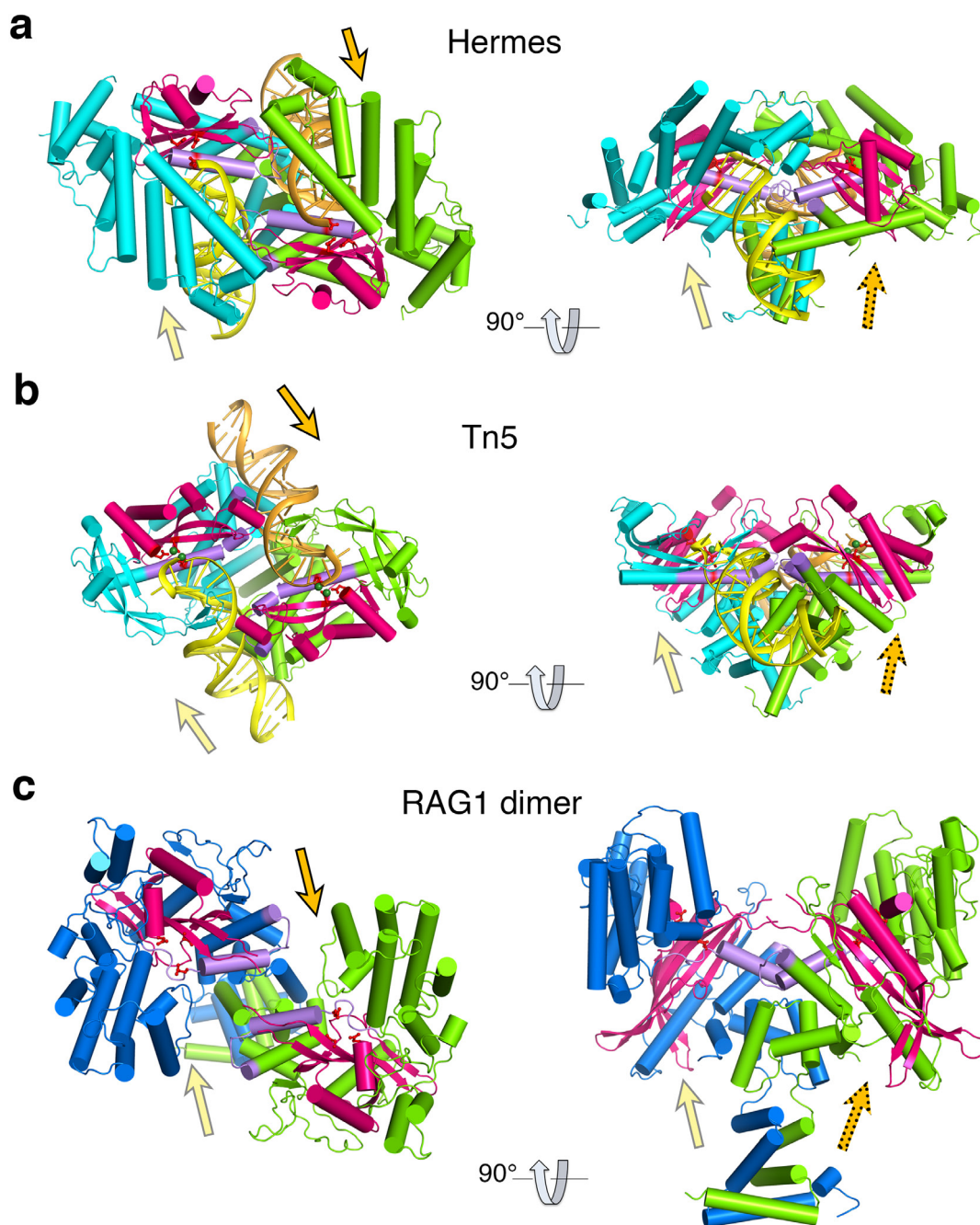
orthogonal views. Each structure is coloured from N to C terminus in blue to red rainbow colours. The loops in RAG2 that interact with RAG1 are labelled. The six β -blades are named by Roman numerals, I–VI, from N to C terminus; four β -strands in each blade are named by Arabic numerals, 1–4.



Extended Data Figure 4 | Comparison of RAG1 and NBD-DNA complex.

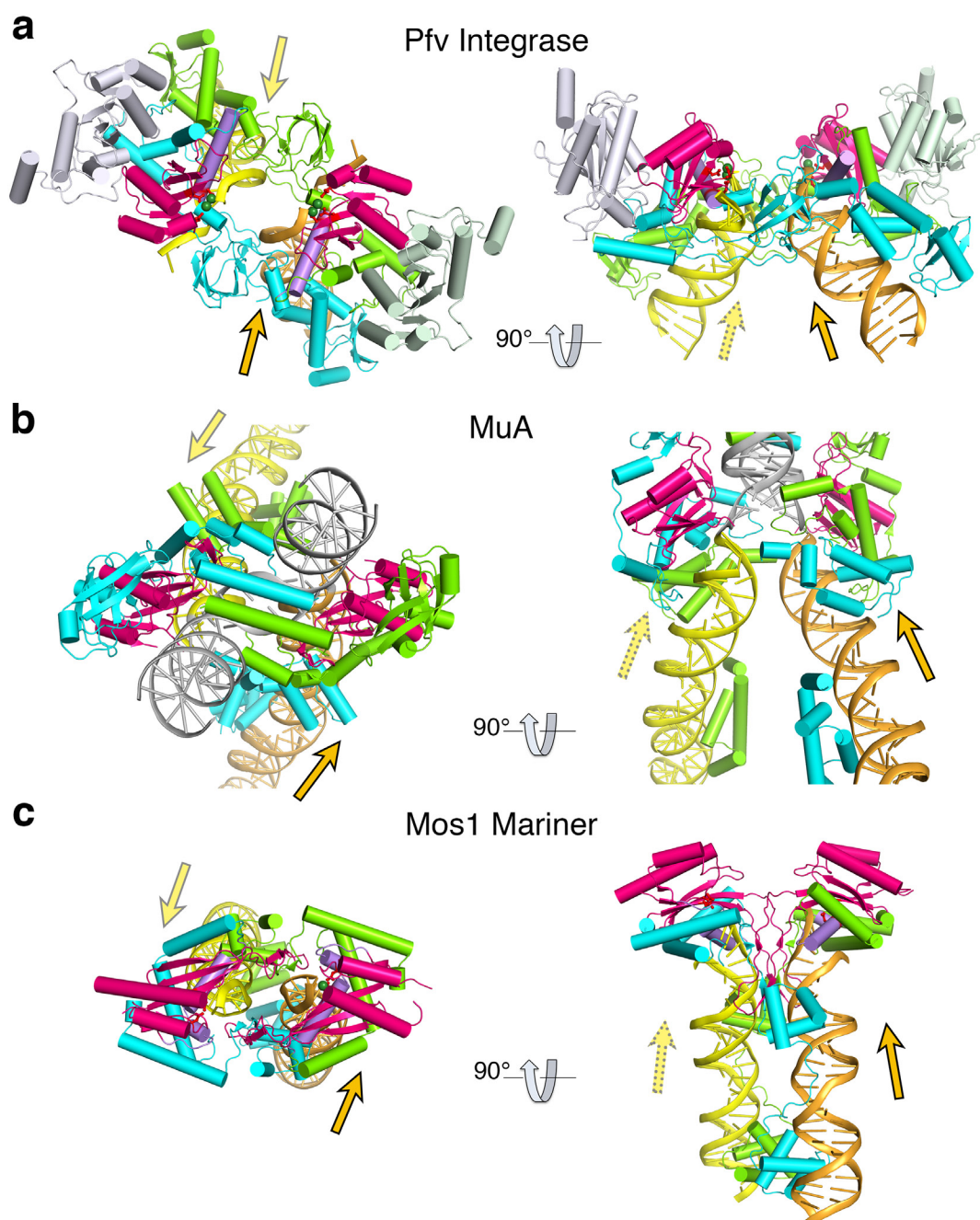
a, The NBD in the RAG1-RAG2 core complex (blue and green) superimposes well with the published structure of the NBD-DNA complex (PDB 3GNA, protein coloured yellow)²². **b**, The twelve SCID/Omenn syndrome mutations in the NBD domain are mapped onto the crystal structure of the NBD-DNA

complex. Six SCID/Omenn syndrome (R391 to R407) mutations are located on a positively charged surface patch that interacts with the nonamer; five remaining SCID/Omenn syndrome mutations (L408 to A441) appear to affect the structural integrity of the NBD, and R446 may interact with the spacer DNA in each RSS.



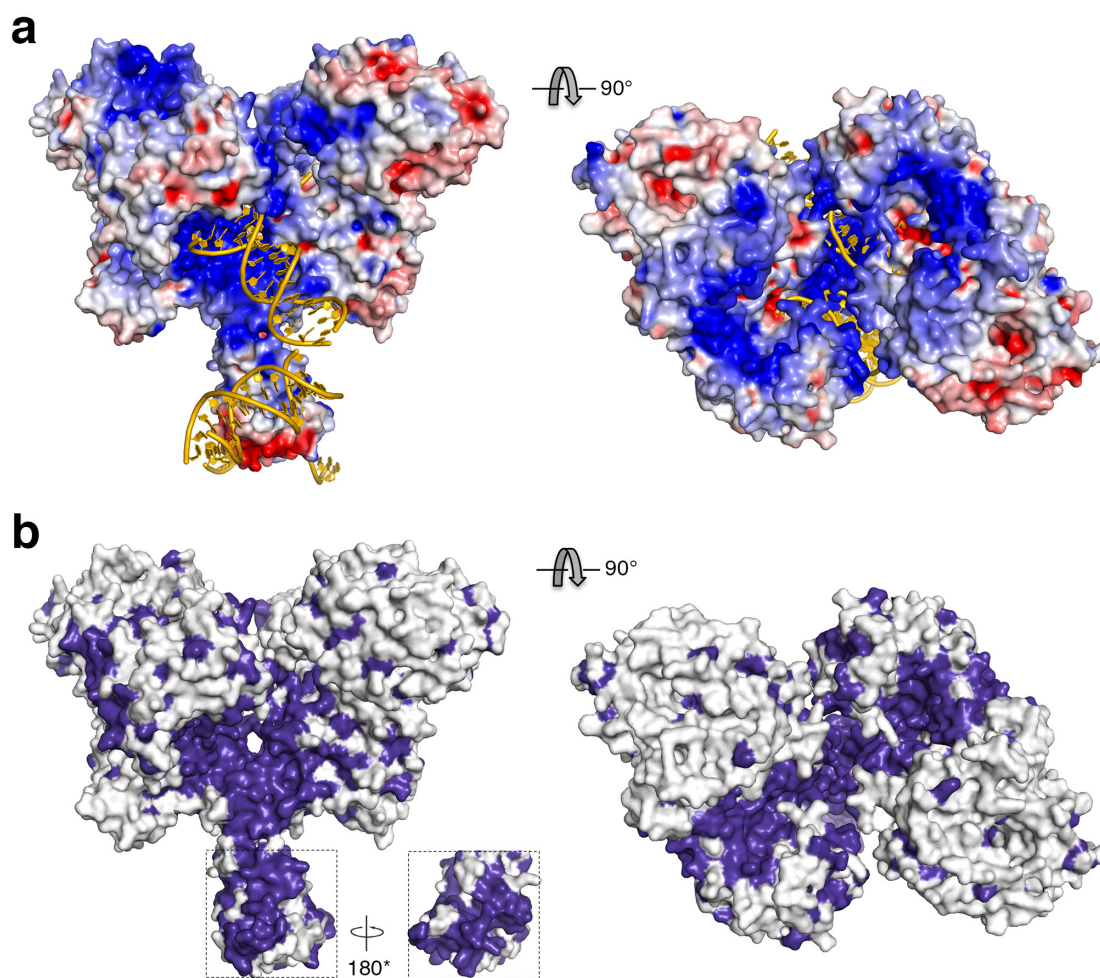
Extended Data Figure 5 | Transposases that form a hairpin intermediate. **a–c**, Hermes (PDB 4D1Q)³⁷ (**a**), bacterial Tn5 (PDB code 1MUS)⁴⁰ (**b**), and RAG1 dimers (**c**) are shown as ribbon diagrams in two orthogonal views, with the dyad perpendicular to the viewing plane (left) or in the plane (right). Each dimer consists of a cyan and a green subunit. The catalytic RNH domains are highlighted in pink, and the conserved catalytic residues are shown as red ball-and-sticks. The catalytic divalent metal ions are shown as green spheres if present. The DNAs, coloured in yellow (cleaved by the cyan subunit) and

orange (cleaved by the green subunit), have similar orientations in the Hermes and Tn5 complexes (as indicated by the arrows). Arrows with dashed outlines indicate that the DNAs are in the back of the viewing plane. Notably, the pair of RNH domains is oriented similarly in all three cases. The predicted orientations of DNAs bound to RAG1 are indicated by the yellow and orange arrows, and the α -helices connected to the third catalytic carboxylates (shown in light purple) probably bridge two DNAs in RAG1 recombinase as in Hermes and Tn5.



Extended Data Figure 6 | Transposases that do not form a hairpin intermediate. **a–c**, Retroviral integrase from Prototype foamy virus (Pfv, PDB 3OS0)³⁶ (**a**), bacterial MuA transposase (PDB code 4FCY)³⁸ (**b**) and eukaryotic Mos1 mariner transposase (PDB 3HOT)³⁹ (**c**) are shown in comparable views and same representations as Hermes, Tn5 and RAG1–RAG2 in Extended Data Fig. 5. Each catalytic dimer consists of a cyan and a green subunit. Two accessory subunits in Pfv are shown in light blue and green, and two accessory subunits of the MuA structure are omitted for clarity. The catalytic RNH

domains are highlighted in pink. The DNAs, coloured in yellow (cleaved by the cyan subunit) and orange (cleaved by the green subunit), have similar orientations (within 30°) as indicated by the arrowheads, but each differs more than 90° from the corresponding DNA in Hermes or Tn5 transposase. The grey DNA in the MuA complex represents the target of transposition. Among these three recombinases, the α -helix that follows the third catalytic carboxylate (coloured in light purple) does not cross over to interact with a second DNA.



Extended Data Figure 7 | Surface potential and conservation of RAG1–RAG2 complex. **a**, Orthogonal views of the electrostatic potential surface of the RAG1–RAG2 structure. Blue indicates positive charges, and red negative. **b**, Orthogonal views of the molecular surface of RAG1–RAG2 with absolutely

conserved residues highlighted in deep purple. The NBD is well conserved. The views with dyad in the plane here are related to the image shown in Fig. 5c by $\sim 50^\circ$ rotation around the dyad.

Extended Data Table 1 | Statistics of native and SeMet SAD data collection and structure refinement

	Native	Crystal #1	Crystal #2	Crystal #3	Crystal #4	Crystal #5	Crystal #6
Space group	C222 ₁	C222 ₁	C222 ₁	C222 ₁	C222 ₁	C222 ₁	C222 ₁
Cell dimensions							
<i>a</i> , <i>b</i> , <i>c</i> (Å)	168.8, 180.1, 200.2	168.7, 179.0, 199.3	168.5, 179.2, 200.3	169.1, 180.1, 200.8	168.7, 180.3, 202.3	169.4, 179.3, 199.7	169.1, 179.6, 200.0
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Absorption (Se)		<i>Peak</i>	<i>Peak</i>	<i>Peak</i>	<i>Peak</i>	<i>Peak</i>	<i>Peak</i>
Wavelength (Å)	1.0000	0.97918	0.97918	0.97918	0.97918	0.97913	0.97918
Resolution [*] (Å)	50-3.2	50.0-3.8	50.0-3.7	50.0-4.0	50.0-3.9	50.0-3.8	50.0-3.9
	(3.31 – 3.2)	(3.94-3.8)	(3.83-3.7)	(4.14-4.0)	(4.04-3.9)	(3.94-3.8)	(4.04-3.9)
<i>R</i> _{merge} [*]	0.105 (0.58)	0.151 (0.778)	0.149 (0.831)	0.199 (0.859)	0.194 (0.796)	0.174 (0.873)	0.200 (0.977)
<i>I</i> / σ <i>I</i> [*]	12.75 (2.23)	17.7 (4.6)	17.5 (3.9)	14.4 (4.1)	14.2 (4.5)	16.8 (4.5)	13.1 (3.8)
Completeness [*] (%)	98.82 (99.9)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Redundancy [*]	7.1 (3.4)	15.0 (15.3)	15.0 (15.3)	14.8 (15.2)	14.9 (15.3)	15.0 (15.3)	15.0 (15.3)
Refinement							
Resolution (Å)	50 - 3.2						
No. reflections	49907						
<i>R</i> _{work} / <i>R</i> _{free}	0.206 / 0.259						
No. atoms							
Protein	14976						
Ligand/ion (Zn ²⁺)	2						
Water	0						
B-factors							
Protein	106.4						
Ligand/ion	85.0						
Water	-						
R.m.s deviations							
Bond lengths (Å)	0.007						
Bond angles (°)	1.114						

Asterisk indicates that data in the highest resolution shell is shown in parenthesis.

Extended Data Table 2 | Missense mutations of RAG1 and RAG2 identified in human SCID/OS patients

Human mutation	Mouse residue	Predicted Structural Effects
RAG1		
R394W/Q	R391A	nonamer binding
R396L/H/C	R393A	nonamer binding
S401P	S398	nonamer binding
T403P	T400	nonamer binding
R404W/A/Q	R401A	nonamer binding
R410Q/W	R407A	nonamer binding
L411P	L408	Structural integrity of NBD
D429G	D426	Structural integrity of NBD
V433M	V430	Structural integrity of NBD
M435V	L432	Structural integrity of NBD
A444V	A441	Structural integrity of NBD
R449K	R446	Probably DNA binding (spacer)
L454Q	L451	Structural integrity of RAG1 dimer
R474S/H/C	R471A	Structure & DNA binding in DDBD
S480G ⁵⁸	S477	Structure & DNA binding in DDBD
L506F	L503	Structural integrity of DDBD
R507W	R504	Solvent exposed, DNA binding?
G516A	G513	Structural integrity of RAG1
W522C	W519	Structural integrity of preR
D539V	D536	Exposed, near RAG1/2 interface
R559S	R556	At the edge of RAG1/2 interface
R561H/C	R558	RAG1/2 interface (T169 of RAG2)
A565D	A562	Structural integrity of preR
S601P	S598	Structural integrity of active site
C602W	C599	Structural integrity of active site
H612R	H609L	Disordered, near RAG2
A622P	A619	Structural integrity of RNH
R624C/H	R621A	Active site, adjacent to D600
E669G	E666	RAG1/2 interface
R699Q/W	R696	Structural stability of RNH
G709D	G706	Structural integrity, active site
R716W	R713A	Structural integrity of RAG12
G720C ⁵⁹	G717	RAG1/2 interface
E722K	E719K	RAG1/2 interface
C730E/F	C727	Structural integrity of ZnC2
L732F/P	L729	Structural integrity of ZnC2
R737H	R734A	Possibly DNA binding (coding end)
H753L	H750A	Structural integrity of RAG1/2
R764P	R761	At the edge of RAG1/2 interface
E770K	E767	RAG1/2 interface
R776Q	R773A	RAG1/2 interface
R778Q/G/W	R775A	Structural integrity of RAG1/2
P786L	P783	RAG1/2 interface
I794T ⁶⁰	I791	Structural integrity of ZnH2
R841W/Q	R838	DNA binding (near heptamer)
N855I	N852	Interacts with 3' end of RSS
L885R	L882	Structural integrity of ZnH2
W896R	W893A	Structural integrity of ZnH2
Y912C	Y909	Structural integrity of ZnH2
I956T	I953R	Structural integrity of ZnH2
R973H/C	R970	Heptamer binding (intra-subunit)
F974L	F971	Structural integrity of CTD
R975W/Q	R972	Structural integrity of CTD
Q981R/P	Q978	Heptamer binding (inter-subunit)
K992E/R	K989	Probable DNA binding
M1006V	M1003	Domain interface of CTD-DDBD
RAG2		
G35V	G35	RAG1/2 interface (E666 of RAG1)
R39G	R39A	RAG1/2 interface (E719, R773)
C41W	C41	Structure, and RAG1/2 interface
G95R	G95	Structure integrity of RAG2
R229E/Q/W	R229	RAG1/2 interface (D546 of RAG1)
M285R	M285	Partially exposed, maybe structure

All SCID/Omenn syndrome mutations here are listed in refs 12 and 13 except for three, for which references are given in the table (refs 58–60). Red residues are the mutations made in mouse⁴³⁴.

Extended Data Table 3 | Mouse RAG1–RAG2 mutations presented in ref. 4

Mutations	Location and potential functional roles
RAG1 mutations: RAG2 binding +, all other -	
K405A/H406A/R407A	Nonamer binding
R748A/H750A	R748 is near ZnC2 and ZnH2 and structurally important. H750 may stabilize the ZnC2 structure.
R773A/R775A	R773 is at RAG1/2 interface between E719 (RAG1) and R39 (RAG2). R775 is exposed to solvent.
H937A/K938A	H937 coordinates the zinc, and K938 forms a salt-bridge with E709 next to the catalytic D708.
H942A	Zinc coordination
R969A/R970A	Next to CTD in the positive groove for DNA binding
RAG1 mutations: RSS binding +, Nicking and Hairpinning -	
K596A	H-bond to the carbonyl of A957 and A955, stabilizes the W956 conformation in the apo-structure
R621A/H	Next to D600 in the active site
R713A	H-bond to “O” of E719, Y725 and I726 (near RAG1/2 interface)
E719K	RAG1/2 interface
R734A	Solvent exposed, but could bind coding-end DNA
W760A	RAG1/2 interface
H795A	In the active site, next to D708
W956A	W956 is near E962 but facing exterior and separate from the active site by the protein backbone.
RAG1 mutations: Nicking +, Hairpinning -	
K608A	Disordered
H609L	Disordered
R855A/K856A	K856 is oriented toward the CTD/DDBD, R855 is solvent exposed
R890A	Near D797 carboxylate, structure integrity, near R855
W893A	Structural integrity of ZnH2
K980A	On the CTD charged surface, potential for heptamer binding
RAG1 mutations: Joining negative or defective	
R401A/R402A	Probable nonamer binding
E423Q	Forming a salt bridge with R407 that probably binds nonamer
R440A	Probable nonamer binding
E547Q	RAG1/2 interface
S723A/C	Adjacent to E719 (OS) and close to R39 of RAG2
RAG1 mutations: Gain-of-function (I2RSS processing)	
E649A	Solvent exposed and adjacent to N961 and S963 near E962.
RAG2 mutations: RAG1 binding+, all other -	
K119A	Solvent exposed and part of the positive top rim of the “Y” shaped RAG1/2 complex
K283A/R	Near M285R (OS), stabilizing the loop of 306-315
RAG2 mutations: RSS binding+, Nicking and Hairpinning -	
K38A/R39A	RAG1/2 interface, possible coding-end binding
RAG2 mutations: Joining negative or defective	
K34A	Near OS mutation G35V, RAG1/2 interface (adjacent to R73 and K97)
K56A/K58A	Solvent exposed, may interact with coding DNA flank
R73A	RAG1/2 interface (adjacent to K34 and K97)
H94A	H94 is a part of the RAG2 structural core and important for the correct folding
K97A	Near OS mutation G95R at RAG1/2 interface (K34 and R73)
K119R	Exposed to solvent and adjacent to K56 and K58
R167A	Near RAG1 interface (N525 of RAG1), forming positive-p stack with W172 (near P674 of RAG1)
361frame shift	without PHD and regulatory domains

Mutations that correspond to SCID/Omenn syndrome mutations in human are highlighted in grey. R795A listed in ref. 4 (highlighted in red) should be H795A.

An ultraluminous quasar with a twelve-billion-solar-mass black hole at redshift 6.30

Xue-Bing Wu^{1,2}, Feige Wang^{1,2}, Xiaohui Fan^{2,3}, Weimin Yi^{4,5,6}, Wenwen Zuo⁷, Fuyan Bian⁸, Linhua Jiang², Ian D. McGreer³, Ran Wang², Jinyi Yang^{1,2}, Qian Yang^{1,2}, David Thompson⁹ & Yuri Beletsky¹⁰

So far, roughly 40 quasars with redshifts greater than $z = 6$ have been discovered^{1–8}. Each quasar contains a black hole with a mass of about one billion solar masses ($10^9 M_\odot$)^{2,6,7,9–13}. The existence of such black holes when the Universe was less than one billion years old presents substantial challenges to theories of the formation and growth of black holes and the coevolution of black holes and galaxies¹⁴. Here we report the discovery of an ultraluminous quasar, SDSS J010013.02+280225.8, at redshift $z = 6.30$. It has an optical and near-infrared luminosity a few times greater than those of previously known $z > 6$ quasars. On the basis of the deep absorption trough¹⁵ on the blue side of the Lyman- α emission line in the spectrum, we estimate the proper size of the ionized proximity zone associated with the quasar to be about 26 million light years, larger than found with other $z > 6.1$ quasars with lower luminosities¹⁶. We estimate (on the basis of a near-infrared spectrum) that the black hole has a mass of $\sim 1.2 \times 10^{10} M_\odot$, which is consistent with the $1.3 \times 10^{10} M_\odot$ derived by assuming an Eddington-limited accretion rate.

High-redshift quasars have been efficiently selected using a combination of optical and near-infrared colours^{3,4}. We have carried out a systematic survey of quasars at $z > 5$ using photometry from the Sloan Digital Sky Survey (SDSS)¹⁷, the two Micron All Sky Survey (2MASS)¹⁸ and the Wide-field Infrared Survey Explorer (WISE)¹⁹, resulting in the discovery of a significant population of luminous high-redshift quasars. SDSS J010013.02+280225.8 (hereafter J0100+2802) was selected as a high-redshift quasar candidate owing to its red optical colour (with SDSS AB magnitudes $i_{AB} = 20.84 \pm 0.06$ and $z_{AB} = 18.33 \pm 0.03$) and a photometric redshift of $z \approx 6.3$. It has bright detections in the 2MASS J, H and K_s bands with Vega magnitudes of 17.00 ± 0.20 , 15.98 ± 0.19 and 15.20 ± 0.16 , respectively; it is also strongly detected in WISE, with Vega magnitudes in W1 to W4 bands of 14.45 ± 0.03 , 13.63 ± 0.03 , 11.71 ± 0.21 and 8.98 ± 0.44 , respectively (see Extended Data Figs 1 and 2 for images in different bands). Its colour in the two bluest WISE bands, W1 and W2, clearly differentiates it from the bulk of stars in our Galaxy²⁰. The object was within the SDSS-III imaging area. It is close to the colour selection boundary of SDSS $z \approx 6$ quasars¹, but was assigned to low priority earlier because of its relatively red $z_{AB} - J$ colour and its bright apparent magnitudes. It is undetected in both radio and X-ray bands by the wide-area, shallow survey instruments.

Initial optical spectroscopy on J0100+2802 was carried out on 29 December 2013 with the Lijiang 2.4-m telescope in China. The low-resolution spectrum clearly shows a sharp break at about 8,800 Å, consistent with a quasar at a redshift beyond 6.2. Two subsequent optical spectroscopic observations were conducted on 9 and 24 January 2014 respectively with the 6.5-m Multiple Mirror Telescope (MMT) and the twin 8.4-m mirror Large Binocular Telescope (LBT) in the USA. The Lyman- α (Ly α) line shown in the spectra confirms that J0100+2802 is a quasar at a redshift of 6.30 ± 0.01 (see Fig. 1 and Methods for details).

We use the multiwavelength photometry to estimate the optical luminosity at rest-frame wavelength 3,000 Å ($L_{3,000}$), which is consistent with that obtained from K-band spectroscopy (see below). The latter gives a more reliable value of $(3.15 \pm 0.47) \times 10^{47} \text{ erg s}^{-1}$, adopting a Λ CDM cosmology with Hubble constant $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, matter density parameter $\Omega_M = 0.30$ and dark energy density parameter $\Omega_\Lambda = 0.7$. Assuming an empirical conversion factor from the luminosity at 3,000 Å to the bolometric luminosity²¹, this gives $L_{\text{bol}} = 5.15 \times L_{3,000} = 1.62 \times 10^{48} \text{ erg s}^{-1} = 4.29 \times 10^{14} L_\odot$ (where L_\odot is the solar luminosity). We obtain a similar result when estimating the bolometric luminosity from the Galactic extinction corrected absolute magnitude at rest-frame 1,450 Å, which is $M_{1450,AB} = -29.26 \pm 0.20$. The luminosity of this

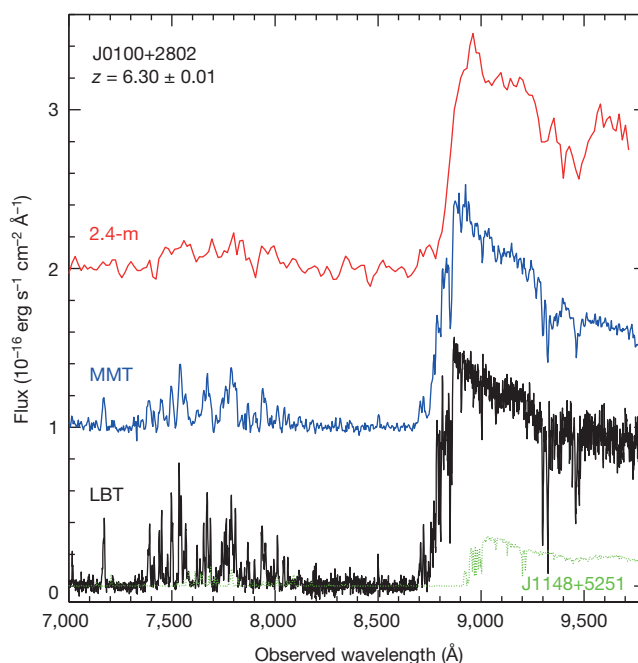


Figure 1 | The optical spectra of J0100+2802. From top to bottom, spectra taken with the Lijiang 2.4-m telescope, the MMT and the LBT (in red, blue and black colours), respectively. For clarity, two spectra are offset upward by one and two vertical units. Although the spectral resolution varies from very low to medium, in all spectra the Ly α emission line, with a rest-frame wavelength of 1,216 Å, is redshifted to around 8,900 Å, giving a redshift of 6.30. J0100+2802 is a weak-line quasar with continuum luminosity about four times higher than that of SDSS J1148+5251 (in green on the same flux scale)¹, which was previously the most luminous high-redshift quasar known at $z = 6.42$.

¹Department of Astronomy, School of Physics, Peking University, Beijing 100871, China. ²Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China. ³Steward Observatory, University of Arizona, Tucson, Arizona 85721-0065, USA. ⁴Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China. ⁵University of Chinese Academy of Sciences, Beijing 100049, China. ⁶Key Laboratory for the Structure and Evolution of Celestial Objects, Chinese Academy of Sciences, Kunming 650011, China. ⁷Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China. ⁸Mount Stromlo Observatory, Research School of Astronomy and Astrophysics, Australian National University, Weston Creek, Australian Capital Territory 2611, Australia. ⁹Large Binocular Telescope Observatory, University of Arizona, Tucson, Arizona 85721, USA. ¹⁰Las Campanas Observatory, Carnegie Institution of Washington, Colina el Pino, Casilla 601, La Serena, Chile.

quasar is roughly 4 times greater than that of the luminous $z = 6.42$ quasar¹ SDSS J1148+5251, and 7 times greater than that of the most distant known quasar⁶ ULAS J1120+0641 ($z = 7.085$); it is the most luminous quasar known at $z > 6$ (see Extended Data Fig. 3).

The rest-frame equivalent width of the Ly α + N V emission lines as measured from the LBT spectrum is roughly 10 Å, suggesting that J0100+2802 is probably a weak-line quasar (WLQ)²². The fraction of WLQs is higher among the $z \approx 6$ quasars compared to those at lower redshift⁸, and a high detection rate of strong millimetre dust continuum in $z \approx 6$ WLQs points to active star formation in these objects²³. Given its extreme luminosity, J0100+2802 will be helpful in the study of the evolutionary stage of WLQs by future (sub)millimetre observations, though the origin of the weak ultraviolet emission line feature of WLQs is still uncertain.

The LBT spectrum of J0100+2802 (Fig. 1) exhibits a deep Gunn–Peterson absorption trough¹⁵ blueward of the Ly α emission. The transmission spectrum (assuming an intrinsic power-law continuum²⁴ of $F_\lambda \propto \lambda^{-1.5}$, where F_λ is the flux density at wavelength λ) is shown in Fig. 2. Complete Gunn–Peterson absorption can also be seen in the Ly α , Ly β and Ly γ transitions. Statistically significant transmission peaks are detected at $z = 5.99$ in both the Ly α and Ly β troughs, and an additional transmission peak is detected at $z = 5.84$ in the Ly β trough. The 2σ lower limit on the Ly α Gunn–Peterson optical depth (τ_α) at $z = 6.00$ – 6.15 is $\tau_\alpha > 5.5$ and the 2σ lower limit for Ly β is $\tau_\beta > 6$, corresponding to an equivalent $\tau_\alpha > 13.5$, following the conversion in literature¹⁶. The characteristics of the intergalactic medium (IGM) transmission along the line of sight of J0100+2802, including the deep Ly α and Ly β troughs, and the narrow, unresolved transmission peaks, are similar to those observed in SDSS J1148+5251, and are consistent with the rapid increase

in the IGM neutral fraction at $z > 5.5$ observed in a large sample of SDSS quasars¹⁶. The size evolution of the quasar proximity zone, which is highly ionized by quasar ultraviolet photons, can also be used to constrain the IGM neutral fraction. The size of the proximity zone is defined by the point where the transmitted flux first drops by a significant amount to below 10% (ignoring small absorption leaks) of the quasar extrapolated continuum emission after the spectrum is smoothed to a resolution of 20 Å (ref. 16). As shown in Fig. 2, J0100+2802 has a much larger proper proximity zone (7.9 ± 0.8 Mpc; 1 Mpc is about 3.26 million light years) than that of other SDSS quasars^{16,25} at $z > 6.1$; its large proximity zone size is expected from the higher level of photo-ionization dominated by quasar radiation.

We obtained the near-infrared J,H,K-band spectra with Gemini and Magellan telescopes on 6 August and 7 October 2014, respectively (see Methods for details). Figure 3 shows the combined optical/near-infrared spectrum of J0100+2802 and the results of fitting the Mg II emission line. The Mg II full-width at half-maximum (FWHM) is $5,130 \pm 150$ km s⁻¹, and the continuum luminosity at the rest-frame wavelength of 3,000 Å is $(3.15 \pm 0.47) \times 10^{47}$ erg s⁻¹. After applying a virial black-hole mass estimator based on the Mg II line²⁶, we estimate its black-hole mass to be $(1.24 \pm 0.19) \times 10^{10} M_\odot$. The uncertainty of black-hole mass does not include the systematic uncertainty of virial black-hole mass estimation, which could be up to a factor of three²⁷. Assuming that this quasar is accreting at the Eddington accretion rate and the bolometric luminosity is close to the Eddington luminosity ($L_{\text{Edd}} = 1.3 \times 10^{38}$ (M/M \odot)), similar to other $z > 6$ quasars¹¹, leads to a black-hole mass of

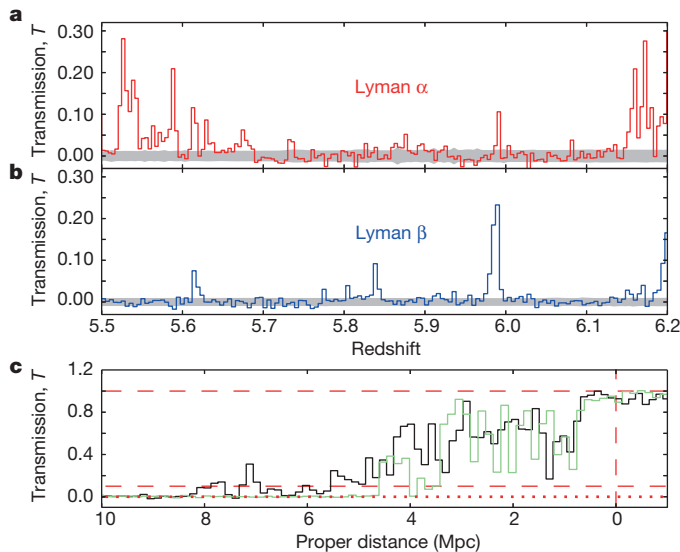


Figure 2 | Transmission in absorption troughs and the proximity zone for J0100+2802. **a, b**, Transmission in Ly α and Ly β absorption troughs (respectively **a**, red; **b**, blue) were calculated by dividing the spectrum by a power-law continuum²⁴, $F_\lambda \propto \lambda^{-1.5}$. The shaded band in both panels shows 1σ standard deviation. The Ly α and Ly β absorption redshifts are given by $\lambda/\lambda_{\text{Ly}\alpha(\text{Ly}\beta)} - 1$, where $\lambda_{\text{Ly}\alpha} = 1,216$ Å and $\lambda_{\text{Ly}\beta} = 1,026$ Å. The optical spectrum exhibits a deep Gunn–Peterson trough and a significant transmission peak at $z = 5.99$. **c**, Transmission in the proximity zone. The proper proximity zone for J0100+2802 (in black) extends to 7.9 ± 0.8 Mpc, a much larger value than those of other $z > 6.1$ quasars, including 4.9 ± 0.6 Mpc for J1148+5251 (in green), consistent with its higher ultraviolet luminosity. The transmission in **c** was calculated by dividing the measured spectrum by a power-law continuum $F_\lambda \propto \lambda^{-1.5}$ plus two Gaussian fittings of Ly α and N V lines. The horizontal dotted line and the two dashed lines denote transmission values of 0, 0.1 and 1.0 respectively, while the vertical dashed line denotes the proper proximity zone size of 0.

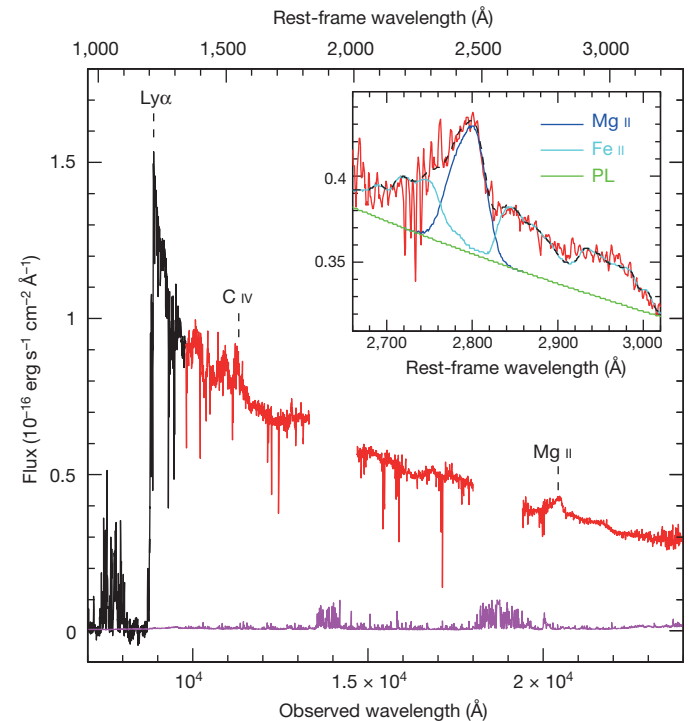


Figure 3 | The combined optical/near-infrared spectrum of J0100+2802 and the fitting of the Mg II line. Main panel, the black line shows the LBT optical spectrum and the red line shows the combined Magellan and Gemini near-infrared J,H,K-band spectra (from left to right, respectively). The gaps between J and H and between H and K bands are ignored due to the low sky transparency there. The magenta line shows the noise spectrum. The main emission lines Ly α , C IV and Mg II are labelled. The details of the absorption lines are described in Extended Data Fig. 4. Inset, fits of the Mg II line (with FWHM of $5,130 \pm 150$ km s⁻¹) and surrounding Fe II emissions. The green, cyan and blue solid lines show the power law (PL), Fe II and Mg II components. The black dashed line shows the sum of these components in comparison with the observed spectrum, denoted by the red line. The black-hole mass is estimated to be $(1.24 \pm 0.19) \times 10^{10} M_\odot$.

$1.3 \times 10^{10} M_{\odot}$ for J0100+2802. Therefore, our observations strongly indicate that J0100+2802 harbours a black hole of mass about $1.2 \times 10^{10} M_{\odot}$, the first such system known at $z > 6$, though black holes of such a size have been found in local giant elliptical galaxies²⁸ and low-redshift quasars²¹.

Although gravitational lensing is a possible explanation for the high luminosity of J0100+2802, we do not expect a large lensing magnification. An LBT K-band image with seeing of $0.4''$ shows a morphology fully consistent with a single point source (Extended Data Fig. 2); and the large size of the quasar proximity zone further supports a high ultraviolet luminosity consistent with the expected photoionization scaling²⁹. However, absorption features at different redshift have been identified from its near-infrared spectroscopy (Extended Data Fig. 4), implying the existence of abundant intervening materials along the line of sight.

J0100+2802 is the only known quasar with a bolometric luminosity higher than $10^{48} \text{ erg s}^{-1}$ and a black-hole mass larger than $5 \times 10^9 M_{\odot}$ at $z \geq 6$. It is also close to being the most luminous quasar with the most massive black hole at any redshift (Fig. 4). The discovery of this single ultraluminous quasar within the entire SDSS footprint ($\sim 13,000$ degrees²) is broadly consistent with the extrapolation of the SDSS $z \approx 6$ quasar luminosity function¹⁶. The number density of such objects would set strong constraints on the early growth of supermassive black holes and the evolution of the high-redshift quasar black-hole mass function^{5,11}. In addition to ULAS J1120+0641 with a $2 \times 10^9 M_{\odot}$ black hole^{6,13} at $z = 7.085$, and a recently discovered $z = 6.889$ quasar with a black hole of $2.1 \times 10^9 M_{\odot}$ (ref. 13), J0100+2802 with a $1.2 \times 10^{10} M_{\odot}$ black hole at $z = 6.30$ presents the next most significant challenge to the Eddington-limited growth of black holes in the early Universe^{11,14}. Its existence also strengthens the claim that supermassive black holes in the early Universe probably grew much more quickly than their host galaxies, as argued from a molecular gas study of $z \approx 6$ quasars³⁰. Therefore, as the most luminous quasar known to date at $z > 6$, J0100+2802 will be a

unique resource for the future study of mass assembly and galaxy formation around the most massive black holes at the end of the epoch of cosmic reionization¹⁶.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 June 2014; accepted 15 January 2015.

1. Fan, X. *et al.* A survey of $z > 5.7$ quasars in the Sloan Digital Sky Survey. II. Discovery of three additional quasars at $z > 6$. *Astron. J.* **125**, 1649–1659 (2003).
2. Jiang, L. *et al.* Gemini near-infrared spectroscopy of luminous $z \sim 6$ quasars: chemical abundances, black hole masses, and Mg II absorption. *Astron. J.* **134**, 1150–1161 (2007).
3. Willott, C. J. *et al.* Four quasars above redshift 6 discovered by the Canada-France high- z quasar survey. *Astron. J.* **134**, 2435–2450 (2007).
4. Jiang, L. *et al.* A survey of $z \sim 6$ quasars in the Sloan Digital Sky Survey deep stripe. I. A flux-limited sample at $z_{AB} < 21$. *Astron. J.* **135**, 1057–1066 (2008).
5. Willott, C. J. *et al.* The Canada-France high- z quasar survey: nine new quasars and the luminosity function at redshift 6. *Astron. J.* **139**, 906–918 (2010).
6. Mortlock, D. J. *et al.* A luminous quasar at a redshift of $z = 7.085$. *Nature* **474**, 616–619 (2011).
7. Venemans, E. P. *et al.* Discovery of three $z > 6.5$ quasars in the VISTA Kilo-Degree Infrared Galaxy (VIKING) Survey. *Astrophys. J.* **779**, 24–36 (2013).
8. Bañados, E. *et al.* Discovery of eight $z \sim 6$ quasars from Pan-STARRS. *Astron. J.* **148**, 14–25 (2014).
9. Willott, C. J., McLure, R. J. & Jarvis, M. J. A $3 \times 10^9 M_{\odot}$ black hole in the quasar SDSS J1148+5251 at $z = 6.41$. *Astrophys. J.* **587**, L15–L18 (2003).
10. Kurk, J. D. *et al.* Black hole masses and enrichment of $z \sim 6$ SDSS quasars. *Astrophys. J.* **669**, 32–44 (2007).
11. Willott, C. J. *et al.* Eddington-limited accretion and the black hole mass function at redshift 6. *Astron. J.* **140**, 546–560 (2010).
12. De Rosa, G. *et al.* Evidence for non-evolving Fe II/Mg II ratios in rapidly accreting $z \sim 6$ QSOs. *Astrophys. J.* **739**, 56–69 (2011).
13. De Rosa, G. *et al.* Black hole mass estimates and emission-line properties of a sample of redshift $z > 6.5$ quasars. *Astrophys. J.* **790**, 145–158 (2014).
14. Volonteri, M. The formation and evolution of massive black holes. *Science* **337**, 544–547 (2012).
15. Gunn, J. E. & Peterson, B. A. On the density of neutral hydrogen in intergalactic space. *Astrophys. J.* **142**, 1633–1641 (1965).
16. Fan, X., Carilli, C. L. & Keating, B. Observational constraints on cosmic reionization. *Annu. Rev. Astron. Astrophys.* **44**, 415–462 (2006).
17. York, D. G. *et al.* The Sloan Digital Sky Survey: technical summary. *Astron. J.* **120**, 1579–1587 (2000).
18. Skrutskie, M. F. *et al.* The Two Micron All Sky Survey (2MASS). *Astron. J.* **131**, 1163–1183 (2006).
19. Wright, E. L. *et al.* The Wide-field Infrared Survey Explorer (WISE): mission description and initial on-orbit performance. *Astron. J.* **140**, 1868–1881 (2010).
20. Wu, X.-B. *et al.* SDSS quasars in the WISE preliminary data release and quasar candidate selection with optical/infrared colours. *Astron. J.* **144**, 49–59 (2012).
21. Shen, Y. *et al.* A catalog of quasar properties from Sloan Digital Sky Survey data release 7. *Astrophys. J.* **194** (suppl.), 45–65 (2011).
22. Fan, X. *et al.* The discovery of a high-redshift quasar without emission lines from Sloan Digital Sky Survey commissioning data. *Astrophys. J.* **526**, L57–L60 (1999).
23. Wang, R. *et al.* Thermal emission from warm dust in the most distant quasars. *Astrophys. J.* **687**, 848–858 (2008).
24. Vanden Berk, D. E. *et al.* Composite quasar spectra from the Sloan Digital Sky Survey. *Astron. J.* **122**, 549–564 (2001).
25. Carilli, C. L. *et al.* Ionization near zones associated with quasars at $z \sim 6$. *Astrophys. J.* **714**, 834–839 (2010).
26. McLure, R. J. & Dunlop, J. S. The cosmological evolution of quasar black hole masses. *Mon. Not. R. Astron. Soc.* **352**, 1390–1404 (2004).
27. Peterson, B. M. *et al.* Central masses and broad-line region sizes of active galactic nuclei. II. A homogeneous analysis of a large reverberation-mapping database. *Astrophys. J.* **613**, 682–699 (2004).
28. McConnell, N. J. *et al.* Two ten-billion-solar-mass black holes at the centres of giant elliptical galaxies. *Nature* **480**, 215–218 (2011).
29. Haiman, Z. *et al.* Constraining reionization with the evolution of the luminosity function of Ly α emitting galaxies. *Astron. J.* **623**, 627–631 (2005).
30. Wang, R. *et al.* Molecular gas in $z \sim 6$ quasar host galaxies. *Astrophys. J.* **714**, 699–712 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements X.-B.W. thanks the NSFC (grant nos 11033001 and 11373008), the Strategic Priority Research Program ‘The Emergence of Cosmological Structures’ of the Chinese Academy of Sciences (grant no. XDB09000000), and the National Key Basic Research Program of China (grant no. 2014CB845700) for support. X.F., R.W. and I.D.M. thank the US NSF (grant nos AST 08-06861 and AST 11-07682) for support. R.W. thanks the NSFC (grant no. 11443002) for support. We acknowledge the support of the staff of the Lijiang 2.4-m telescope. Funding for the telescope was provided by the Chinese Academy of Sciences and the People’s Government of Yunnan Province. This

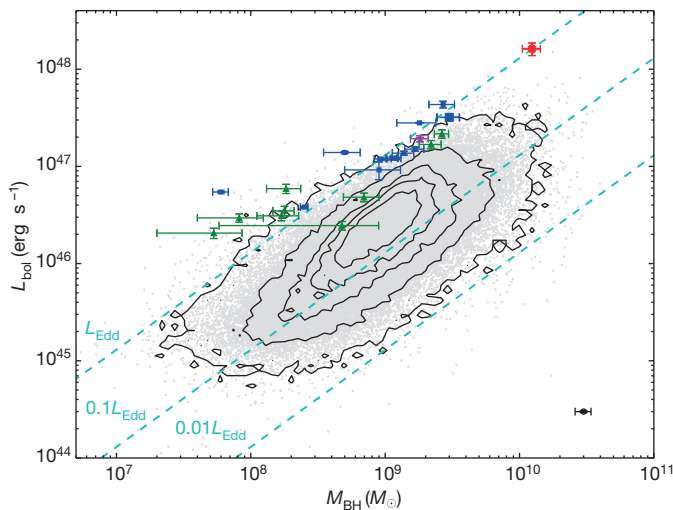


Figure 4 | Distribution of quasar bolometric luminosities, L_{bol} , and black-hole masses, M_{BH} , estimated from the Mg II lines. The red circle at top right represents J0100+2802. The small blue squares denote SDSS high-redshift quasars^{2,10,12}, and the large blue square represents J1148+5251. The green triangles denote CFHQS high-redshift quasars^{11,12}. The purple star denotes ULAS J1120+0641 at $z = 7.085$ (ref. 6). Black contours (which indicate 1σ to 5σ significance from inner to outer) and grey dots denote SDSS low-redshift quasars²¹ (with broad absorption line quasars excluded). Error bars represent the 1σ standard deviation, and the mean error bar for low-redshift quasars is presented in the bottom-right corner. The dashed lines denote the luminosity in different fractions of the Eddington luminosity, L_{Edd} . Note that the black-hole mass and bolometric luminosity are calculated using the same method and the same cosmology model as in the present Letter, and the systematic uncertainties (not included in the error bars) of virial black-hole masses could be up to a factor of three²⁷.

research uses data obtained through the Telescope Access Program (TAP), which has been funded by the Strategic Priority Research Program 'The Emergence of Cosmological Structures' (grant no. XDB09000000), National Astronomical Observatories, Chinese Academy of Sciences, and the Special Fund for Astronomy from the Ministry of Finance of China. We thank D. Osip for help with Magellan/FIRE spectroscopy, and Y.-L. Ai, L. C. Ho, Y. Shen and J.-G. Wang for suggestions about data analyses. We acknowledge the use of SDSS, 2MASS and WISE data, and of the MMT, LBT, Gemini and Magellan telescopes; detailed acknowledgments of these facilities can be found in Supplementary Information.

Author Contributions X.-B.W., F.W. and X.F. planned the study, and wrote the draft version of the paper. All other co-authors contributed extensively and equally to the observations, data analyses and writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to X.-B.W. (wuxb@pku.edu.cn).

METHODS

The optical spectroscopy on J0100+2802 was first carried out on 29 December 2013 with the Yunnan Fainter Object Spectrograph and Camera (YFOSC) of the Lijiang 2.4-m telescope in China. We used a very low resolution grism (G12, at a dispersion of 18 Å per pixel) and took 3,000 s exposure on this target. The spectrum clearly shows a sharp break at about 8,800 Å and no significant emissions blueward, consistent with a quasar spectrum at a redshift beyond 6.2. To confirm this discovery, two subsequent optical spectroscopic observations were obtained on 9 and 24 January 2014 with the 6.5-m Multiple Mirror Telescope (MMT) and the twin 8.4-m mirror Large Binocular Telescope (LBT) in the USA, respectively. The low to medium resolution spectra, obtained with 1,200 s exposure using the MMT Red Channel (at a dispersion of 3.6 Å per pixel) and 2,400 s exposure with the LBT Multi-Object Double CCD Spectrographs/Imagers (MODS)³¹ (at a dispersion of 1.8 Å per pixel) respectively, explicitly confirm that SDSS J0100+2802 is a quasar at redshift 6.30 ± 0.01 (obtained by the Ly α line).

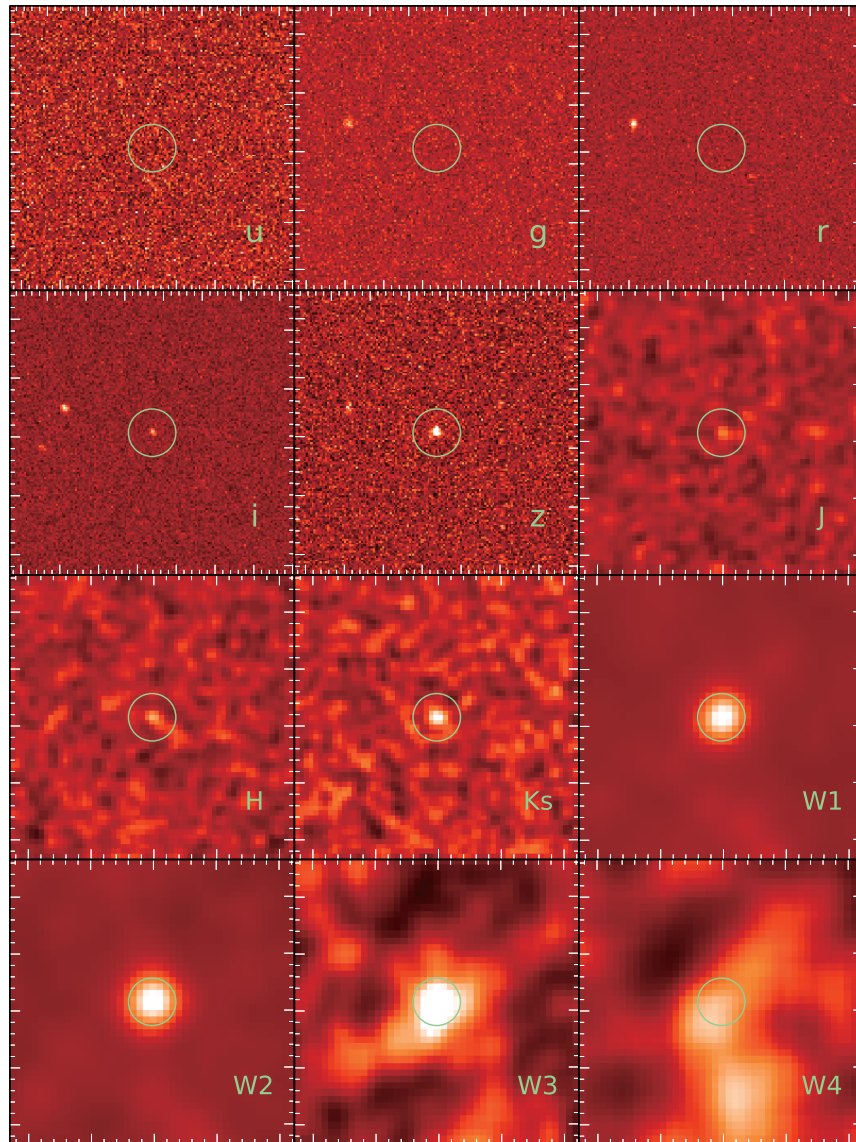
The near-infrared K-band spectroscopy on J0100+2802 was carried out with LBT/LUCI-1 on 2 January 2014. Owing to the short exposure time (15 min), the spectrum is of modest signal-to-noise ratio (S/N). Although the Mg II line was clearly detected, the noisy LBT spectrum did not allow us to accurately measure the line width. To improve the quality of the near-infrared spectrum, we obtained J,H,K-band spectroscopy with Gemini/GNIRS and Magellan/FIRE on 6 August and 7 October 2014, respectively. The exposure time was 3,600 s for GNIRS and 3,635 s for FIRE. The FIRE spectrum has higher S/N (about 30 in K band) and higher spectral resolution ($R = \lambda/\Delta\lambda \approx 6,000$) than the GNIRS spectrum (with S/N of about 10 in K band and $R \approx 1,800$). In order to achieve the best spectral quality, we combined the FIRE and the GNIRS spectra, and scaled the combined spectrum according to its 2MASS J,H,K_s-band magnitudes. The Mg II line shown in the K-band spectrum gives the same redshift as that given by the Ly α line in the optical spectrum. The high-quality J,H,K-band spectra also clearly display abundant absorption features, which have been identified as being from intervening or associated systems with redshifts from 2.33 to 6.14 (Extended Data Fig. 4).

After redshift and Galactic extinction corrections, the rest-frame H- and K-band spectrum is decomposed into a pseudo-continuum and the Mg II emission line. The pseudo-continuum consists of a power-law continuum and Fe II emissions, and

is fitted within the rest-frame wavelength range between 2,000 Å and 3,200 Å by excluding the boundary region between H and K bands where the sky transparency is lower. An Fe II template^{32,33} is adopted for the fitting of Fe II emissions. The Mg II emission line is fitted with two broad Gaussian components. The four Mg II absorption lines near the redder part of the Mg II emission line are also fitted as four Gaussian lines in order to remove their effects on the fittings of Mg II and Fe II emission lines. The overall FWHM of the Mg II emission line is $\sim 5,130 \text{ km s}^{-1}$ with an uncertainty of 150 km s^{-1} . The continuum has a slope of -1.43 and the continuum luminosity at the rest-frame wavelength of 3,000 Å ($L_{3,000}$) is $(3.15 \pm 0.47) \times 10^{47} \text{ erg s}^{-1}$. The Fe II to Mg II line ratio is 2.56 ± 0.18 , which is consistent with the mean value of other $z > 6$ quasars^{12,13}. After applying a virial black-hole mass estimator based on the Mg II line²⁶, we estimate its black-hole mass to be $(1.24 \pm 0.19) \times 10^{10} M_{\odot}$. Although the systematic uncertainty of virial black-hole mass estimation can be up a factor of three²⁷, our result still strongly indicates that J0100+2802 hosts a central black hole with mass close to $1.2 \times 10^{10} M_{\odot}$. This is also well consistent with a black-hole mass obtained by assuming an Eddington luminosity of J0100+2802, which leads to a mass of $1.3 \times 10^{10} M_{\odot}$. Considering the contribution of Balmer continuum, as done for other $z > 6$ quasars^{12,13}, leads to a decrease of $L_{3,000}$ to $(2.90 \pm 0.44) \times 10^{47} \text{ erg s}^{-1}$, an increase of FWHM of Mg II to $5,300 \pm 200 \text{ km s}^{-1}$ and yields a black-hole mass of $(1.26 \pm 0.21) \times 10^{10} M_{\odot}$. Therefore, the effect of considering Balmer continuum is insignificant for the black-hole mass measurement of J0100+2802. In addition, if we adopt a different virial black-hole mass scaling relation³⁴, the black-hole mass changes to $(1.07 \pm 0.14) \times 10^{10} M_{\odot}$, which is still well consistent with the result we obtained above.

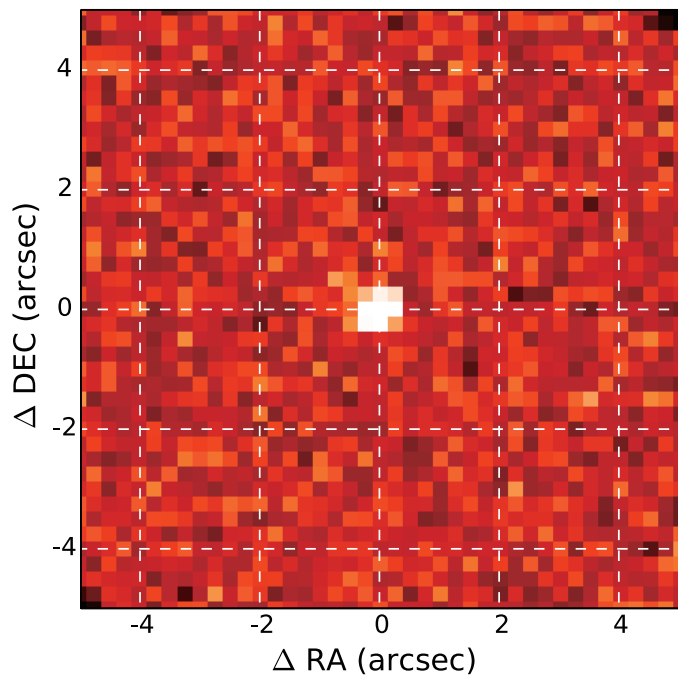
Sample size. No statistical methods were used to predetermine sample size.

31. Pogge, R. W. *et al.* The multi-object double spectrographs for the Large Binocular Telescope. *Proc. SPIE* **7735**, 9–16 (2010).
32. Vestergaard, M. & Wilkes, B. J. An empirical ultraviolet template for iron emission in quasars as derived from I Zwicky 1. *Astrophys. J. Suppl. Ser.* **134**, 1–33 (2001).
33. Salviander, S. *et al.* The black hole mass-galaxy bulge relationship for QSOs in the Sloan Digital Sky Survey data release 3. *Astrophys. J.* **662**, 131–144 (2007).
34. Vestergaard, M. & Osmer, P. S. Mass functions of the active black holes in distant quasars from the Large Bright Quasar Survey, the Bright Quasar Survey, and the color-selected sample of the SDSS Fall Equatorial Stripe. *Astrophys. J.* **699**, 800–816 (2009).

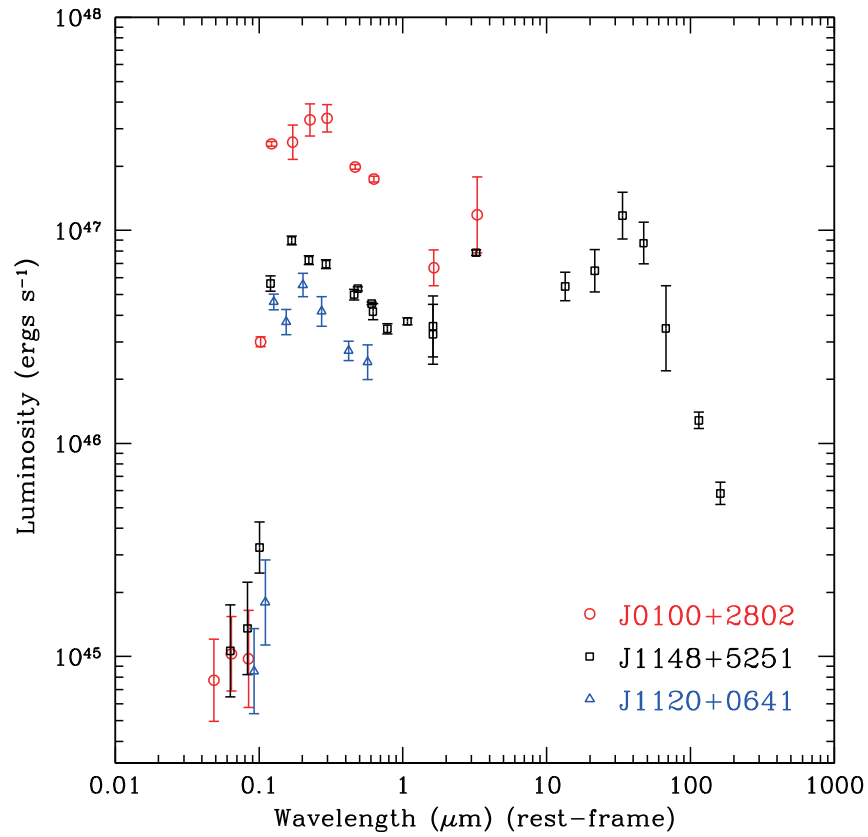


Extended Data Figure 1 | Images of J0100+2802 in SDSS, 2MASS and WISE bands. J0100+2802 is undetected in SDSS u,g,r bands (top row) but is relatively bright in other bands (lower three rows). It is consistent with a point

source in the bands with high signal-to-noise detections. The size is $1' \times 1'$ for all images. The green circle represents an angular size of $10''$ in each image.

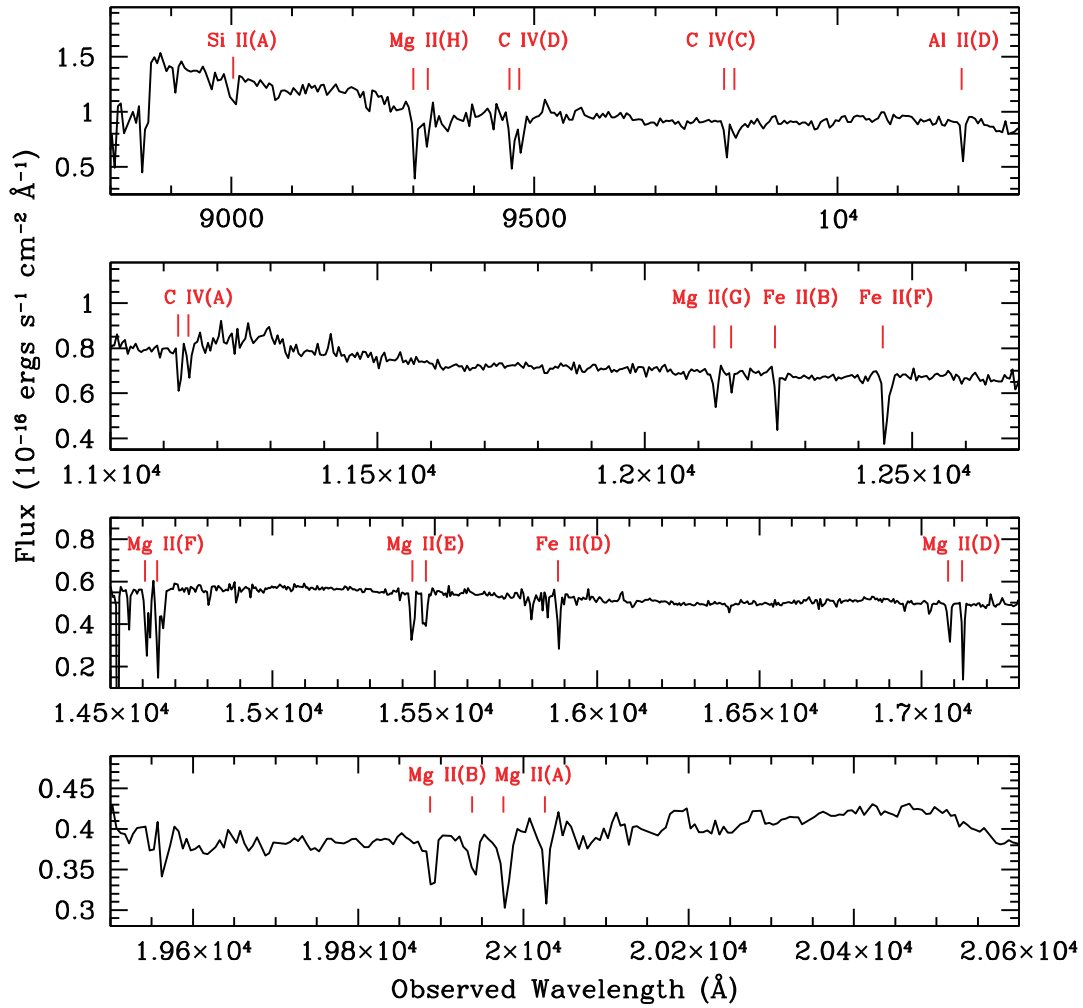


Extended Data Figure 2 | The LBT K-band image of J0100+2802. The size is $10'' \times 10''$. The horizontal and vertical axes denote the offsets in right ascension (ΔRA) and in declination ($\Delta Dec.$). The image, with seeing of $0.4''$, shows a morphology fully consistent with a point source.



Extended Data Figure 3 | The rest-frame spectral energy distributions of J0100+2802, J1148+5251 and ULAS J1120+0641. The redshifts of these three quasars are 6.30, 6.42 and 7.085, respectively. The luminosity of J0100+2802 in the ultraviolet/optical bands is about four times higher than that

of J1148+5251, and seven times higher than that of ULAS J1120+0641. The photometric data are from literature for J1148+5251 and J1120+0641. The error bars show the 1σ standard deviation.



Extended Data Figure 4 | The major absorption features identified from optical and near-infrared spectroscopy of J0100+2802. Most of them are from Mg II, C IV and Fe II. The labels from A to H correspond to the redshifts of

absorption materials at 6.14, 6.11, 5.32, 5.11, 4.52, 4.22, 3.34 and 2.33, respectively. Studies of intervening and associated absorption systems will be discussed elsewhere.

Quantum teleportation of multiple degrees of freedom of a single photon

Xi-Lin Wang^{1,2}, Xin-Dong Cai^{1,2}, Zu-En Su^{1,2}, Ming-Cheng Chen^{1,2}, Dian Wu^{1,2}, Li Li^{1,2}, Nai-Le Liu^{1,2}, Chao-Yang Lu^{1,2} & Jian-Wei Pan^{1,2}

Quantum teleportation¹ provides a ‘disembodied’ way to transfer quantum states from one object to another at a distant location, assisted by previously shared entangled states and a classical communication channel. As well as being of fundamental interest, teleportation has been recognized as an important element in long-distance quantum communication², distributed quantum networks³ and measurement-based quantum computation^{4,5}. There have been numerous demonstrations of teleportation in different physical systems such as photons^{6–8}, atoms⁹, ions^{10,11}, electrons¹² and superconducting circuits¹³. All the previous experiments were limited to the teleportation of one degree of freedom only. However, a single quantum particle can naturally possess various degrees of freedom—internal and external—and with coherent coupling among them. A fundamental open challenge is to teleport multiple degrees of freedom simultaneously, which is necessary to describe a quantum particle fully and, therefore, to teleport it intact. Here we demonstrate quantum teleportation of the composite quantum states of a single photon encoded in both spin and orbital angular momentum. We use photon pairs entangled in both degrees of freedom (that is, hyper-entangled) as the quantum channel for teleportation, and develop a method to project and discriminate hyper-entangled Bell states by exploiting probabilistic quantum non-demolition measurement, which can be extended to more degrees of freedom. We verify the teleportation for both spin–orbit product states and hybrid entangled states, and achieve a teleportation fidelity ranging from 0.57 to 0.68, above the classical limit. Our work is a step towards the teleportation of more complex quantum systems, and demonstrates an increase in our technical control of scalable quantum technologies.

Quantum teleportation is a linear operation applied to quantum states, and so teleporting multiple degrees of freedom (DoFs) should be possible in theory¹. Suppose Alice wishes to teleport to Bob the composite quantum state of a single photon (photon 1; Fig. 1a), encoded in both the spin angular momentum (SAM) and the orbital angular momentum (OAM) as follows:

$$|\phi\rangle_1 = \alpha|0\rangle_1^s|0\rangle_1^o + \beta|0\rangle_1^s|1\rangle_1^o + \gamma|1\rangle_1^s|0\rangle_1^o + \delta|1\rangle_1^s|1\rangle_1^o$$

Here $|0\rangle^s$ and $|1\rangle^s$ denote horizontal and vertical polarizations of the SAM; $|0\rangle^o$ and $|1\rangle^o$ refer to right-handed and left-handed OAMs of $+\hbar$ and $-\hbar$, respectively; and α , β , γ and δ are complex numbers satisfying $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. For Alice to do so, she and Bob first need to share a hyper-entangled photon pair (photons 2 and 3), which is simultaneously entangled in both SAM and OAM:

$$\begin{aligned} |\xi\rangle_{23} &= |\phi^-\rangle_{23}|\omega^+\rangle_{23} \\ &= \frac{1}{2}(|0\rangle_2^s|0\rangle_3^s - |1\rangle_2^s|1\rangle_3^s)(|0\rangle_2^o|0\rangle_3^o + |1\rangle_2^o|1\rangle_3^o) \end{aligned}$$

Here we use $|\phi^\pm\rangle = (|0\rangle^s|0\rangle^s \pm |1\rangle^s|1\rangle^s)/\sqrt{2}$ and $|\psi^\pm\rangle = (|0\rangle^s|1\rangle^s \pm |1\rangle^s|0\rangle^s)/\sqrt{2}$ to denote the four Bell states encoded in SAM, and $|\omega^\pm\rangle = (|0\rangle^o|0\rangle^o \pm |1\rangle^o|1\rangle^o)/\sqrt{2}$ and $|\chi^\pm\rangle = (|0\rangle^o|1\rangle^o \pm |1\rangle^o|0\rangle^o)/\sqrt{2}$ to denote the four Bell states in OAM. Their tensor products result in a group of 16 hyper-entangled Bell states.

A crucial step in the teleportation is to perform a two-particle joint measurement of photons 1 and 2, projecting them onto the basis of the

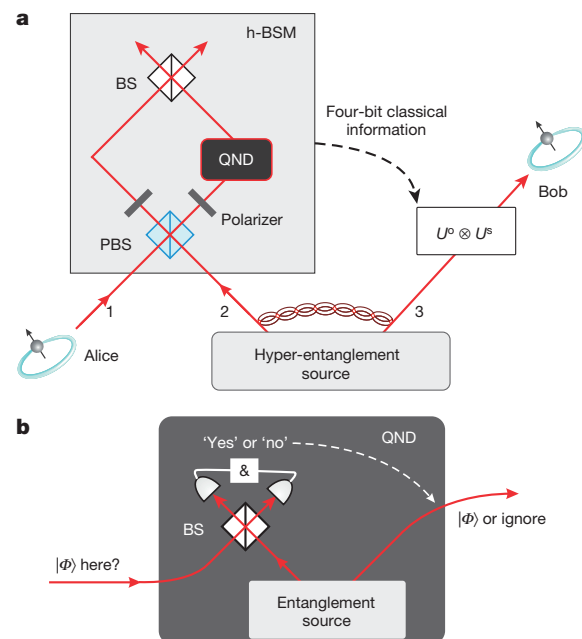


Figure 1 | Scheme for quantum teleportation of the spin–orbit composite states of a single photon. **a**, Alice wishes to teleport to Bob the quantum state of a single photon 1 encoded in both its SAM and OAM. To do so, Alice and Bob need to share a hyper-entangled photon pair 2–3. Alice then carries out an h-BSM assisted by a QND measurement (see main text for details) and sends the results as four-bit classical information to Bob. On receiving Alice’s h-BSM result, Bob can apply appropriate Pauli operations (denoted U^o and U^s for the OAM and SAM DoFs, respectively) on photon 3 to convert it into the original state of photon 1. The active feed-forward is essential for a full, deterministic teleportation. In our present proof-of-principle experiment, we did not apply feed-forward but used post-selection to verify the success of teleportation. BS, beam splitter. **b**, Teleportation-based probabilistic QND measurement with an ancillary entangled photon pair. An incoming photon can cause a coincidence detection after the beam splitter, which heralds its presence and meanwhile fully teleports its arbitrary unknown quantum state ($|\phi\rangle$) to a free-flying photon. In the case of no incoming photon, the coincidence event after the beam splitter cannot happen, thus indicating the photon’s absence.

¹Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China. ²CAS Centre for Excellence and Synergetic Innovation Centre in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China.

16 orthogonal and complete hyper-entangled Bell states, and discriminating one of them; for example

$$\begin{aligned} |\xi\rangle_{12} &= |\phi^-\rangle_{12}|\omega^+\rangle_{12} \\ &= \frac{1}{2}(|0\rangle_1^s|0\rangle_2^s - |1\rangle_1^s|1\rangle_2^s)(|0\rangle_1^o|0\rangle_2^o + |1\rangle_1^o|1\rangle_2^o) \end{aligned}$$

This process is referred to as hyper-entangled Bell state measurement (h-BSM). After the h-BSM that projects photons 1 and 2 onto the state $|\xi\rangle_{12}$, photon 3 will be projected onto the initial state of photon 1:

$$|\varphi\rangle_3 = \alpha|0\rangle_3^s|0\rangle_3^o + \beta|0\rangle_3^s|1\rangle_3^o + \gamma|1\rangle_3^s|0\rangle_3^o + \delta|1\rangle_3^s|1\rangle_3^o$$

With equal probabilities of 1/16, photons 1 and 2 can also be projected onto one of the other 15 hyper-entangled Bell states (Methods). The h-BSM results can be broadcast as four-bit classical information, which will allow Bob to apply appropriate Pauli operations to perfectly reconstruct the initial composite state of photon 1.

Experimental realization of the above teleportation protocol poses significant challenges to the coherent control of multiple particles and multiple DoFs simultaneously. The most difficult task is to implement the h-BSM, because it would normally require coherently controlled gates between independent quantum bits (qubits) of different DoFs. Moreover, with multiple DoFs, it is necessary to measure one DoF without disturbing any other. With linear operations only, previous theoretical work¹⁴ has suggested that it is impossible to discriminate the hyper-entangled states unambiguously. This challenge has been overcome in our work.

Figure 1 illustrates our linear optical scheme for teleporting the spin-orbit composite state. The h-BSM is implemented in a step-by-step manner, as a combination of two separate BSMs. First, photons 1 and 2 are sent through a polarizing beam splitter (PBS), an optical device that transmits horizontal polarizations ($|0\rangle^s$) and reflects vertical polarization ($|1\rangle^s$). After the PBS, we post-select the event that there is one and only one photon in each output. Such an event can occur only if the two input photons have the same SAM (both are transmitted ($|0\rangle_1^s|0\rangle_2^s$) or reflected ($|1\rangle_1^s|1\rangle_2^s$)), which projects the SAM part of the wavefunction into the two-dimensional subspace spanned by $|\phi^\pm\rangle = (|0\rangle^s|0\rangle^s \pm |1\rangle^s|1\rangle^s)/\sqrt{2}$. At both outputs of the PBS, we add two polarizers, projecting the two photons into the diagonal basis $(|0\rangle^s + |1\rangle^s)/\sqrt{2}$. It should be noted that the PBS is not OAM-preserving¹⁵, because the reflection at the PBS flips the sign of the OAM qubit; that is, $|0\rangle_1^s|0\rangle_1^o \rightarrow |0\rangle_1^s|0\rangle_1^o$, $|0\rangle_1^s|1\rangle_1^o \rightarrow |0\rangle_1^s|1\rangle_1^o$, $|1\rangle_1^s|0\rangle_1^o \rightarrow i|1\rangle_2^s|1\rangle_2^o$ and $|1\rangle_1^s|1\rangle_1^o \rightarrow i|1\rangle_2^s|0\rangle_2^o$. Thus, both the SAM and the OAM must be taken into account as in a molecular-like coupled state. A detailed mathematical treatment is presented in Methods, showing that the PBS and two polarizers select the four following states out of the total 16 hyper-entangled Bell states: $|\phi^+\rangle_{12}|\omega^-\rangle_{12}$, $|\phi^-\rangle_{12}|\omega^+\rangle_{12}$, $|\phi^-\rangle_{12}|\chi^+\rangle_{12}$ and $|\phi^-\rangle_{12}|\chi^-\rangle_{12}$.

Second, having measured and filtered out the SAM, we perform BSM on the remaining OAM qubit. The two single photons that emerge from the PBS are superposed on a beam splitter (Fig. 1a). Only the asymmetric Bell state will lead to a coincidence detection where there is one and only one photon in each output¹⁶, whereas for the three other symmetric Bell states, the two input photons will coalesce to a single output mode. Note again that reflection at the beam splitter inverts the OAM sign. Therefore, the state $|\omega^-\rangle_{12}$ can be distinguished by a coincidence detection in separate outputs, and $|\omega^+\rangle_{12}$ can be discriminated by measuring two orthogonal OAMs in either output. In total, these two steps would allow an unambiguous discrimination of the two hyper-entangled Bell states $|\phi^+\rangle_{12}|\omega^-\rangle_{12}$ and $|\phi^-\rangle_{12}|\omega^+\rangle_{12}$.

However, these interferometric processes cannot be simply cascaded. They depend on the assumption that two input photons are from different paths. When directly connected, the 50% chance of both photons coalescing into a single output spatial mode after the first ‘interferometer’ would remain undetected (no ‘coincidence’ event occurs), and would further induce erroneous detection of a coincidence event after the second

interferometer, thus failing both BSMs. This is a ubiquitous and important problem in linear optical quantum information processing^{5,17}.

Our remedy is to exploit quantum non-demolition (QND) measurement, whereby a single photon is observed without destroying it and keeping its quantum information intact. Interestingly, quantum teleportation itself can be used for probabilistic QND detection^{7,18}. As shown in Fig. 1b, another pair of photons entangled in OAM is used as ancillary. The procedure is a standard teleportation. If there is an incoming photon, a two-photon coincidence detection behind a beam splitter can occur with 50% efficiency, triggering a successful BSM, which heralds the presence of the incoming photon and teleports the full quantum state of the incoming photon to a freely propagating photon. If there is no incoming photon, two-photon coincidence behind the beam splitter cannot occur, in which case we will know and will ignore the outgoing photon. Thus, using the QND measurement, the BSM interferometers can be concatenated. We note that using a QND in one of the arm of the beam splitter is sufficient owing to the conserved total number of eventually registered photons. Our protocol can identify two hyper-entangled Bell states with an overall efficiency of 1/32 (Methods). The QND method can also be used to boost the efficiency of photonic quantum logic gates in the Knill–Laflamme–Milburn scheme for scalable optical quantum computing⁵.

Figure 2 shows the experimental set-up for quantum teleportation of the spin-orbit composite state of a single photon. Passing a femtosecond-pulsed laser through three type-I β -barium borate crystals generates three photon pairs^{19–21}, engineered in different forms (Methods). The first photon pair (1–t) is used to prepare a heralded single photon (1) to be teleported, triggered by the detection of its sister photon (t). The second pair (2–3) is created in the hyper-entangled state $|\xi\rangle_{23}$. The third pair (4–5) is prepared in the ancillary OAM-entangled state $|\omega^+\rangle_{45}$ for the teleportation-based QND measurement.

We prepare five different initial states to be teleported: $|\varphi\rangle_A = |0\rangle^s|0\rangle^o$, $|\varphi\rangle_B = |1\rangle^s|1\rangle^o$, $|\varphi\rangle_C = (|0\rangle^s + |1\rangle^s)(|0\rangle^o + |1\rangle^o)/2$, $|\varphi\rangle_D = (|0\rangle^s + i|1\rangle^s)(|0\rangle^o + i|1\rangle^o)/2$ and $|\varphi\rangle_E = (|0\rangle^s|0\rangle^o + |1\rangle^s|1\rangle^o)/\sqrt{2}$. These states can be grouped into three categories: $|\varphi\rangle_A$ and $|\varphi\rangle_B$ are product states of the two DoFs in the computational basis; $|\varphi\rangle_C$ and $|\varphi\rangle_D$ are products states of the two DoFs in the superposition basis; and $|\varphi\rangle_E$ is a spin-orbit hybrid entangled state. The four product states are prepared by independent single-qubit rotations, using wave plates for SAM and spiral phase plates (SPPs) or binary phase plates (BPPs) for OAM. The entangled state is generated by counter-propagating a SPP inside a Sagnac interferometer (Methods).

The implementation of the h-BSM and the QND measurement requires Hong–Ou–Mandel-type interference²² between indistinguishable single photons with good time, spatial and spectral overlap. The two photons are synchronized to arrive at the PBS and beam splitters within 10 fs of each other, a delay that is much smaller than the coherence time of the down-converted photons (~ 448 fs), which is stretched by narrowband spectral filtering (~ 3 nm). A step-by-step verification of the two-photon interference as a function of temporal delay is presented in Methods and Extended Data Fig. 1. We observe a visibility of 0.75 ± 0.03 for the interference of the two SAM-encoded photons at the PBS, and visibilities of 0.73 ± 0.03 and 0.69 ± 0.03 for the interferences of two OAM-encoded photons at beam splitter 1 and, respectively, beam splitter 2.

The final verification of the teleportation results relies on the coincidence detection counts of the six photons encoded in both DoFs, which would suffer from a low rate. It should be noted that all the previous experiments^{15,20,21,23–26} with OAM states have never gone beyond two single photons. We overcome this technical challenge by preparing high-brightness, hyper-entangled photons and designing dual-channel and high-efficiency OAM measurement devices (Methods).

To evaluate the performance of the teleportation operation, we measure the fidelity of the teleported state, $F = \text{Tr}(\hat{\rho}|\varphi\rangle\langle\varphi|)$, which is defined as the overlap of the ideal teleported state ($|\varphi\rangle$) and the measured density matrix ($\hat{\rho}$). Conditioned on the detection of the trigger photon

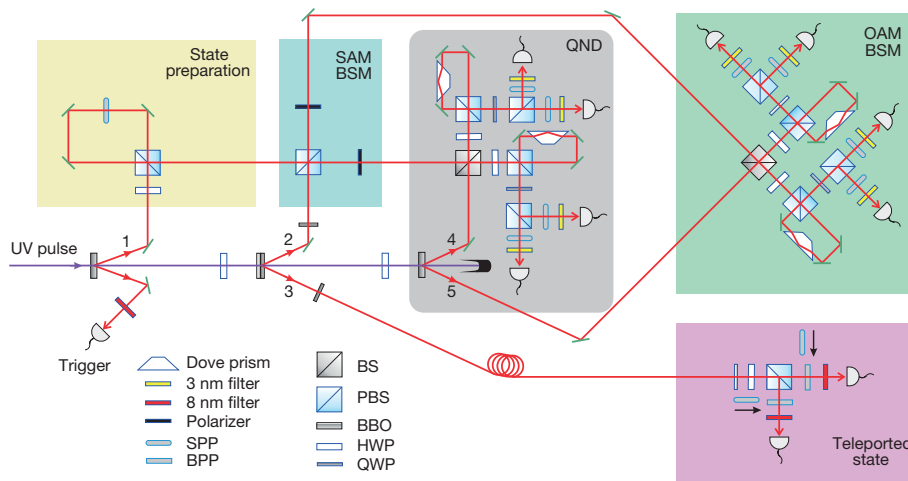


Figure 2 | Experimental set-up for teleporting multiple properties of a single photon. A pulsed ultraviolet (UV) laser is focused on three β -barium borate (BBO) crystals and produces three photon pairs in spatial modes 1–t, 2–3 and 4–5. Triggered by its sister photon, t, photon 1 is initialised in various spin-orbit composite states ($|\varphi\rangle_{A,\dots,E}$) to be teleported. The second pair, 2–3, is hyper-entangled in both SAM and OAM. The third pair, 4–5, is OAM-entangled. The h-BSMs for photons 1 and 2 are performed in three steps: (1) SAM BSM; (2) QND measurement; (3) OAM BSM. The teleported state is measured separately in SAM and OAM: a PBS, a half-wave plate (HWP) and a quarter-wave plate (QWP) are combined for SAM qubit analysis, and an SPP or a BPP together with a single-mode fibre are used for OAM qubit analysis.

t and the four-photon coincidence after the h-BSM, we register the photon counts of teleported photon 3 and analyse its composite state. The fidelity measurements for the product states are straightforward because we can measure the SAM and OAM qubits separately. We measure the final state in the $(|0\rangle^s, |1\rangle^s)$ and $(|0\rangle^o, |1\rangle^o)$ bases for the teleportation of $|\varphi\rangle_A$ and $|\varphi\rangle_B$, in the $(|0\rangle^s \pm |1\rangle^s)/\sqrt{2}$ and $(|0\rangle^o \pm |1\rangle^o)/\sqrt{2}$ bases for the teleportation of $|\varphi\rangle_C$, and in the $(|0\rangle^s \pm i|1\rangle^s)/\sqrt{2}$ and $(|0\rangle^o \pm i|1\rangle^o)/\sqrt{2}$ bases for the teleportation of $|\varphi\rangle_D$. The data in Fig. 3 yield teleportation fidelities of 0.68 ± 0.04 , 0.66 ± 0.04 , 0.62 ± 0.04 and 0.63 ± 0.04 for $|\varphi\rangle_A$, $|\varphi\rangle_B$, $|\varphi\rangle_C$ and, respectively, $|\varphi\rangle_D$.

The fidelity of the entangled state $|\varphi\rangle_E$, where the SAM and OAM qubits are not separable, can be decomposed as $F_E = \text{Tr}[\hat{\rho}(I + \hat{\sigma}_x^s \hat{\sigma}_x^o - \hat{\sigma}_y^s \hat{\sigma}_y^o + \hat{\sigma}_z^s \hat{\sigma}_z^o)]/4$, where σ_x , σ_y and σ_z are the Pauli operators. The expectation values of the joint observables $\hat{\sigma}_x^s \hat{\sigma}_x^o$, $\hat{\sigma}_y^s \hat{\sigma}_y^o$ and $\hat{\sigma}_z^s \hat{\sigma}_z^o$ can be obtained by local measurements in the corresponding basis for the two DoFs, $(|0\rangle \pm |1\rangle)/\sqrt{2}$, $(|0\rangle \pm i|1\rangle)/\sqrt{2}$ and $(|0\rangle, |1\rangle)$, respectively. Our experimental results on the three different bases are presented in Fig. 4a–c, from which we determine a teleportation fidelity of 0.57 ± 0.02 for the entangled state.

We note that all reported data are without background subtraction. The main sources of error include double pair emission, imperfection in the initial states, entanglement of photons 2–3 and 4–5, two-photon interference and OAM measurement. We note that the teleportation fidelities of the states in the three categories are affected differently by errors from different sources (Methods). Despite the experimental noise, the measured fidelities (summarized in Fig. 4d) of the five teleported states are all well above 0.40—the classical limit, defined as the optimal state-estimation fidelity on a single copy of a two-qubit system²⁷. These results prove the successful realization of quantum teleportation of the spin-orbit composite state of a single photon. Furthermore, for the entangled state, $|\varphi\rangle_E$, we emphasize that the teleportation fidelity exceeds the threshold of 0.5 required to prove the presence of entanglement²⁸, which demonstrates that the hybrid entanglement of different DoFs inside a quantum particle can survive teleportation.

We have reported the quantum teleportation of multiple properties of a single quantum particle, demonstrating the ability to coherently control and simultaneously teleport a single object with multiple DoFs that forms in a hybrid entangled state. It is interesting to note that these DoFs can also be in a fully undefined state, such as being part of a hyper-entangled pair, which would lead to the protocol of hyper-entanglement swapping²⁹. Our methods can be generalized to more DoFs (see Methods for a universal scheme). The efficiency of teleportation, which is limited mainly by the efficiency of h-BSM (1/32 in the present experiment), can be enhanced by using more ancillary photons, quantum encoding, embedded teleportation tricks, high-efficiency single-photon detectors and active feed-forward, in a spirit similar to the Knill–Laflamme–Milburn scheme⁵. We did not implement the feed-forward in the current

experiment, but it could be done using electro-optical modulators for both the SAM and the OAM qubits (see Methods for a detailed protocol). Although the present work is based on linear optics and single photons, the multi-DoF teleportation protocol is by no means limited to this system, but can also be applied to other quantum systems such as trapped electrons¹², atoms⁹ and ions^{10,11}.

As well as being of fundamental interest, the methods developed in this work on the manipulation of quantum states of multiple DoFs will

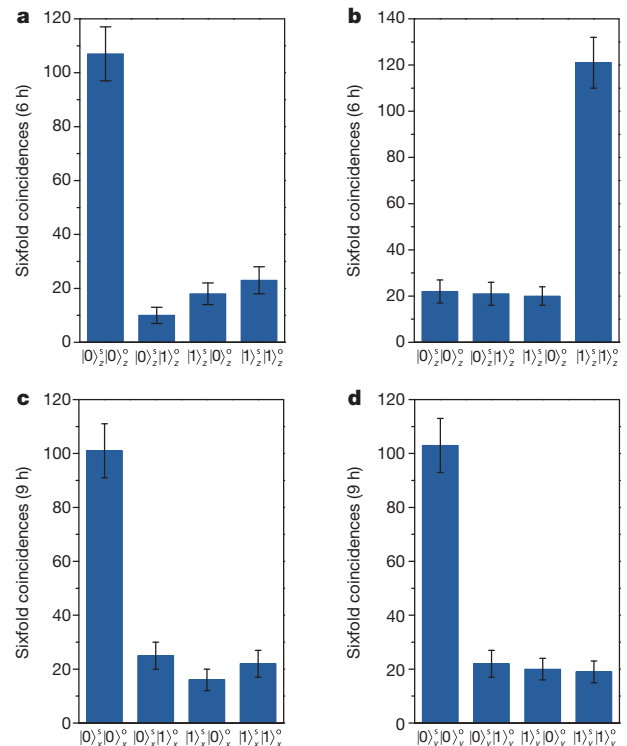


Figure 3 | Experimental results for quantum teleportation of spin-orbit product states $|\varphi\rangle_A$, $|\varphi\rangle_B$, $|\varphi\rangle_C$ and $|\varphi\rangle_D$ of a single photon. **a, b,** Measurement results of the final state of the teleported photon 3 in the $(|0\rangle^s, |1\rangle^s)$ and $(|0\rangle^o, |1\rangle^o)$ bases for the $|\varphi\rangle_A$ (**a**) and $|\varphi\rangle_B$ (**b**) teleportation experiment. **c,** The results for the $|\varphi\rangle_C$ teleportation, measured in the $(|0\rangle^s \pm |1\rangle^s)/\sqrt{2}$ and $(|0\rangle^o \pm |1\rangle^o)/\sqrt{2}$ bases. **d,** The results for the $|\varphi\rangle_D$ teleportation, measured in the $(|0\rangle^s \pm i|1\rangle^s)/\sqrt{2}$ and $(|0\rangle^o \pm i|1\rangle^o)/\sqrt{2}$ bases. The x axis uses Pauli notation for both DoFs: $(|0\rangle_z, |1\rangle_z) = (|0\rangle, |1\rangle)$, $(|0\rangle_x, |1\rangle_x) = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $(|0\rangle_y, |1\rangle_y) = (|0\rangle \pm i|1\rangle)/\sqrt{2}$. The y axis is six-photon coincidence counts. Error bars, 1 s.d., calculated from Poissonian counting statistics of the raw detection events.

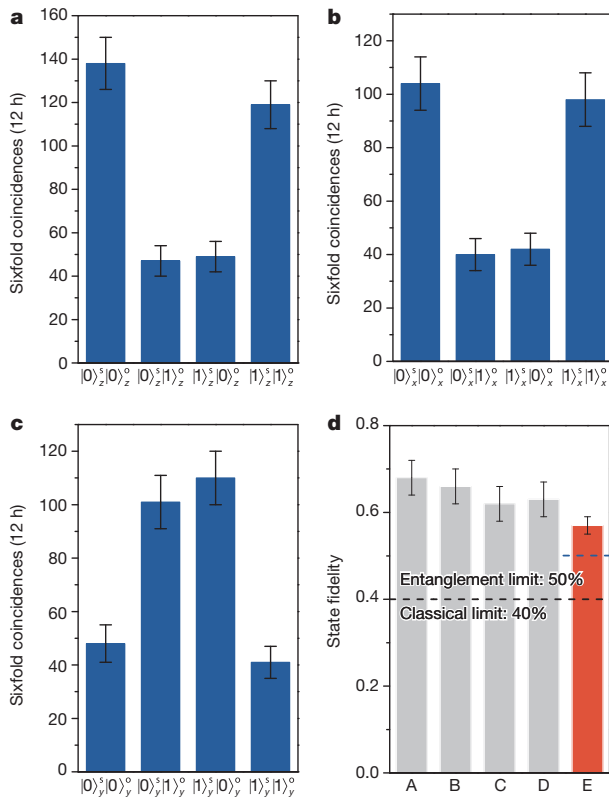


Figure 4 | Experimental results for quantum teleportation of spin-orbit entanglement of a single photon. **a–c**, To determine the state fidelity of the teleported entangled state, $|\varphi\rangle_E = (|0\rangle_s^s|0\rangle_x^o + |1\rangle_s^s|1\rangle_x^o)/\sqrt{2}$, three measurement bases are required: $(|0\rangle_s^s, |1\rangle_x^o)$ and $(|0\rangle_s^s, |1\rangle_y^o)$ (**a**), $(|0\rangle_s^s \pm |1\rangle_s^s)/\sqrt{2}$ and $(|0\rangle_x^o \pm |1\rangle_x^o)/\sqrt{2}$ (**b**), and $(|0\rangle_s^s \pm |1\rangle_s^s)/\sqrt{2}$ and $(|0\rangle_x^o \pm |1\rangle_y^o)/\sqrt{2}$ (**c**). These are used to extract the expectation values of the joint Pauli observables $\hat{\sigma}_x^s \hat{\sigma}_x^o$, $\hat{\sigma}_x^s \hat{\sigma}_y^o$ and $\hat{\sigma}_y^s \hat{\sigma}_y^o$, respectively. Each measurement takes 12 h. The data set in **a** determines the population of the two desired terms: $|0\rangle_s^s|0\rangle_x^o$ and $|1\rangle_s^s|1\rangle_x^o$ in the entangled state, $|\varphi\rangle_E$. The data set in **b** and **c** measured in the superposition basis determines the coherence of the entangled state. We use the same Pauli notation as in Fig. 3. **d**, A summary of the teleportation fidelities for the states $|\varphi\rangle_{A, \dots, E}$. Error bars, 1 s.d., deduced from propagated Poissonian counting statistics of the raw detection events.

open up new possibilities in quantum technologies. Controlling multiple DoFs makes complete SAM Bell-state analysis²³ and alignment-free quantum communication³⁰ possible, for example. Carrying multiple DoFs on photons can enhance the information capacity in quantum communication protocols such as quantum super-dense coding²³. Moreover, combining the entanglement of multiple (M) photons and high (N) dimensions from the multiple DoFs would allow the generation of hyper-entanglement with an expanded Hilbert space that grows in size as N^M , which would provide a versatile platform in the near future for demonstrations of complex quantum communication and quantum computing protocols, as well as extreme violations of Bell inequalities.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 September 2014; accepted 15 January 2015.

1. Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classic and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).

2. Briegel, H. J., Dur, W., Cirac, J. I. & Zoller, P. Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998).
3. Kimble, H. J. The quantum internet. *Nature* **453**, 1023–1030 (2008).
4. Gottesman, D. & Chuang, I. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature* **402**, 390–393 (1999).
5. Knill, E., Laflamme, R. & Milburn, G. J. A scheme for efficient quantum computation with linear optics. *Nature* **409**, 46–52 (2001).
6. Bouwmeester, D. *et al.* Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
7. Marcikic, I., de Riedmatten, H., Tittel, W., Zbinden, H. & Gisin, N. Long-distance teleportation of qubits at telecommunication wavelengths. *Nature* **421**, 509–513 (2003).
8. Takeda, S., Mizuta, T., Fuwa, M., van Loock, P. & Furusawa, A. Deterministic quantum teleportation of photonic quantum bits by a hybrid technique. *Nature* **500**, 315–318 (2013).
9. Krauter, H. *et al.* Deterministic quantum teleportation between distant atomic objects. *Nature Phys.* **9**, 400–404 (2013).
10. Riebe, M. *et al.* Deterministic quantum teleportation with atoms. *Nature* **429**, 734–737 (2004).
11. Barrett, M. D. *et al.* Deterministic quantum teleportation of atomic qubits. *Nature* **429**, 737–739 (2004).
12. Pfaff, W. *et al.* Unconditional quantum teleportation between distant solid-state quantum bits. *Science* **345**, 532–535 (2014).
13. Steffen, L. *et al.* Deterministic quantum teleportation with feed-forward in a solid state system. *Nature* **500**, 319–322 (2013).
14. Wei, T. C., Barreiro, J. T. & Kwiat, P. G. Hyperentangled Bell-state analysis. *Phys. Rev. A* **75**, 060305(R) (2007).
15. Nagali, E. *et al.* Optimal quantum cloning of orbital angular momentum photon qubits through Hong–Ou–Mandel coalescence. *Nature Photon.* **3**, 720–723 (2009).
16. Weinfurter, H. Experimental Bell-state analysis. *Europhys. Lett.* **25**, 559–564 (1994).
17. Pan, J.-W. *et al.* Multiphoton entanglement and interferometry. *Rev. Mod. Phys.* **84**, 777–838 (2012).
18. Jacobs, B. C., Pittman, T. B. & Franson, J. D. Quantum relays and noise suppression using linear optics. *Phys. Rev. A* **66**, 052307 (2002).
19. Kwiat, P. G. *et al.* Ultrabright source of polarization-entangled photons. *Phys. Rev. A* **60**, 773–776 (1999).
20. Barreiro, J. T., Langford, N. K., Peters, N. A. & Kwiat, P. G. Generation of hyper-entangled photon pairs. *Phys. Rev. Lett.* **95**, 260501 (2005).
21. Mair, A., Vaziri, A., Weihs, G. & Zeilinger, A. Entanglement of the orbital angular momentum states of photons. *Nature* **412**, 313–316 (2001).
22. Hong, C. K., Ou, Z. Y. & Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
23. Barreiro, J. T., Wei, T.-C. & Kwiat, P. G. Beating the channel capacity limit for linear photonic superdense coding. *Nature Phys.* **4**, 282–286 (2008).
24. Nagali, E. *et al.* Quantum information transfer from spin to orbital angular momentum of photons. *Phys. Rev. Lett.* **103**, 013601 (2009).
25. Leach, J. *et al.* Quantum correlations in optical angle-orbital angular momentum variables. *Science* **329**, 662–665 (2010).
26. Fickler, R. *et al.* Quantum entanglement of high angular momenta. *Science* **338**, 640–643 (2012).
27. Hayashi, A., Hashimoto, T. & Horibe, M. Reexamination of optimal quantum state estimation of pure states. *Phys. Rev. A* **72**, 032325 (2005).
28. Gühne, O. & Toth, G. Entanglement detection. *Phys. Rep.* **474**, 1–75 (2009).
29. Żukowski, M., Zeilinger, A., Horne, M. A. & Ekert, A. “Event-ready-detectors” Bell experiment via entanglement swapping. *Phys. Rev. Lett.* **71**, 4287–4290 (1993).
30. D’Ambrosio, V. *et al.* Complete experimental toolbox for alignment-free quantum communication. *Nature Commun.* **3**, 961 (2012).

Acknowledgements This work was supported by the National Natural Science Foundation of China, the Chinese Academy of Sciences and the National Fundamental Research Program (grant no. 2011CB921300).

Author Contributions C.-Y.L. and J.-W.P. had the idea for and designed the research; X.-L.W., X.-D.C., Z.-E.S., M.-C.C., D.W., L.L., N.-L.L. and C.-Y.L. performed the experiment; X.-L.W., M.-C.C., C.-Y.L. and J.-W.P. analysed the data; C.-Y.L. and J.-W.P. wrote the paper with input from all authors; and N.-L.L., C.-Y.L. and J.-W.P. supervised the whole project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.-Y.L. (cyu@ustc.edu.cn) or J.-W.P. (pan@ustc.edu.cn).

METHODS

The protocol for teleporting a spin-orbit composite quantum state. The combined state of photons 1, 2 and 3 can be rewritten in the basis of the 16 orthogonal and complete hyper-entangled Bell states as follows:

$$\begin{aligned}
 |\varphi\rangle_1|\xi\rangle_{23} = & \frac{1}{4}[(|\phi^+\rangle_{12}|\omega^+\rangle_{12})\hat{\sigma}_{3z}^s|\varphi\rangle_3 \\
 & + (|\phi^+\rangle_{12}|\omega^-\rangle_{12})\hat{\sigma}_{3z}^s\hat{\sigma}_{3z}^o|\varphi\rangle_3 \\
 & + (|\phi^+\rangle_{12}|\chi^+\rangle_{12})\hat{\sigma}_{3z}^s\hat{\sigma}_{3x}^o|\varphi\rangle_3 \\
 & - (|\phi^+\rangle_{12}|\chi^-\rangle_{12})i\hat{\sigma}_{3z}^s\hat{\sigma}_{3y}^o|\varphi\rangle_3 \\
 & + (|\phi^-\rangle_{12}|\omega^+\rangle_{12})|\varphi\rangle_3 \\
 & + (|\phi^-\rangle_{12}|\omega^-\rangle_{12})\hat{\sigma}_{3z}^o|\varphi\rangle_3 \\
 & + (|\phi^-\rangle_{12}|\chi^+\rangle_{12})\hat{\sigma}_{3x}^o|\varphi\rangle_3 \\
 & - (|\phi^-\rangle_{12}|\chi^-\rangle_{12})i\hat{\sigma}_{3y}^o|\varphi\rangle_3 \\
 & + (|\psi^+\rangle_{12}|\omega^+\rangle_{12})i\hat{\sigma}_{3y}^s|\varphi\rangle_3 \\
 & + (|\psi^+\rangle_{12}|\omega^-\rangle_{12})i\hat{\sigma}_{3y}^s\hat{\sigma}_{3z}^o|\varphi\rangle_3 \\
 & + (|\psi^+\rangle_{12}|\chi^+\rangle_{12})i\hat{\sigma}_{3y}^s\hat{\sigma}_{3x}^o|\varphi\rangle_3 \\
 & + (|\psi^+\rangle_{12}|\chi^-\rangle_{12})\hat{\sigma}_{3y}^s\hat{\sigma}_{3y}^o|\varphi\rangle_3 \\
 & - (|\psi^-\rangle_{12}|\omega^+\rangle_{12})\hat{\sigma}_{3x}^s|\varphi\rangle_3 \\
 & - (|\psi^-\rangle_{12}|\omega^-\rangle_{12})\hat{\sigma}_{3x}^s\hat{\sigma}_{3z}^o|\varphi\rangle_3 \\
 & - (|\psi^-\rangle_{12}|\chi^+\rangle_{12})\hat{\sigma}_{3x}^s\hat{\sigma}_{3x}^o|\varphi\rangle_3 \\
 & + (|\psi^-\rangle_{12}|\chi^-\rangle_{12})i\hat{\sigma}_{3x}^s\hat{\sigma}_{3y}^o|\varphi\rangle_3]
 \end{aligned}$$

where the triples $(\hat{\sigma}_x^s, \hat{\sigma}_y^s, \hat{\sigma}_z^s)$ and $(\hat{\sigma}_x^o, \hat{\sigma}_y^o, \hat{\sigma}_z^o)$ are the Pauli operators for the SAM and, respectively, the OAM qubits. It indicates that, regardless of the unknown state $|\varphi\rangle_1$, the 16 measurement outcomes are equally likely, each with a probability of 1/16. By carrying out the hyper-entangled Bell state measurement (h-BSM) on photons 1 and 2 to unambiguously distinguish one from the group of 16 hyper-entangled Bell states, Alice can project photon 3 onto one of the 16 corresponding states. After Alice tells Bob her h-BSM result as four-bit classical information via a classical communication channel, Bob can convert the state of his photon 3 into the original state by applying appropriate two-qubit local unitary transformations $[(\hat{I}, \hat{\sigma}_x^s, \hat{\sigma}_y^s, \hat{\sigma}_z^s) \otimes (\hat{I}, \hat{\sigma}_x^o, \hat{\sigma}_y^o, \hat{\sigma}_z^o)]$.

Two-photon interference of spin-orbit composite state on a PBS. The input and output of photons 1 and 2 encoded in SAM and OAM on a PBS as shown in Fig. 1 can be summarized as follows: $|0\rangle_1^s|0\rangle_1^o \rightarrow |0\rangle_1^s|0\rangle_1^o$, $|0\rangle_1^s|1\rangle_1^o \rightarrow |0\rangle_1^s|1\rangle_1^o$, $|1\rangle_1^s|0\rangle_1^o \rightarrow i|1\rangle_2^s|1\rangle_2^o$, $|1\rangle_1^s|1\rangle_1^o \rightarrow i|1\rangle_2^s|0\rangle_2^o$. Note that the PBS is not OAM-preserving, because the reflection flips the sign of OAM. Therefore, the output state for each of the input 16 hyper-entangled Bell states can be listed below:

$$|\phi^-\rangle_{12}|\omega^+\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'D_2' + A_1'A_2')|\omega^+\rangle_{1'2'} \quad (1)$$

$$|\phi^+\rangle_{12}|\omega^-\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'D_2' + A_1'A_2')|\omega^-\rangle_{1'2'} \quad (2)$$

$$|\phi^-\rangle_{12}|\chi^+\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'D_2' + A_1'A_2')|\chi^+\rangle_{1'2'} \quad (3)$$

$$|\phi^-\rangle_{12}|\chi^-\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'D_2' + A_1'A_2')|\chi^-\rangle_{1'2'} \quad (4)$$

$$|\phi^+\rangle_{12}|\omega^+\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'A_2' + A_1'D_2')|\omega^+\rangle_{1'2'} \quad (5)$$

$$|\phi^-\rangle_{12}|\omega^-\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'A_2' + A_1'D_2')|\omega^-\rangle_{1'2'} \quad (6)$$

$$|\phi^+\rangle_{12}|\chi^+\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'A_2' + A_1'D_2')|\chi^+\rangle_{1'2'} \quad (7)$$

$$|\phi^+\rangle_{12}|\chi^-\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(D_1'A_2' + A_1'D_2')|\chi^-\rangle_{1'2'} \quad (8)$$

$$\begin{aligned}
 |\psi^+\rangle_{12}|\omega^+\rangle_{12} \rightarrow & \frac{i}{2}(D_1'r_1'D_1'l_1' - A_1'r_1'A_1'l_1' \\
 & + D_2'r_2'D_2'l_2' - A_2'r_2'A_2'l_2')
 \end{aligned} \quad (9)$$

$$\begin{aligned}
 |\psi^-\rangle_{12}|\omega^+\rangle_{12} \rightarrow & \frac{i}{2}(D_1'r_1'D_1'l_1' - A_1'r_1'A_1'l_1' \\
 & - D_2'r_2'D_2'l_2' + A_2'r_2'A_2'l_2')
 \end{aligned} \quad (10)$$

$$\begin{aligned}
 |\psi^+\rangle_{12}|\omega^-\rangle_{12} \rightarrow & \frac{i}{2}(D_1'r_1'A_1'l_1' - A_1'r_1'D_1'l_1' \\
 & + D_2'r_2'A_2'l_2' - A_2'r_2'D_2'l_2')
 \end{aligned} \quad (11)$$

$$\begin{aligned}
 |\psi^-\rangle_{12}|\omega^-\rangle_{12} \rightarrow & \frac{i}{2}(D_1'r_1'A_1'l_1' - A_1'r_1'D_1'l_1' \\
 & - D_2'r_2'A_2'l_2' + A_2'r_2'D_2'l_2')
 \end{aligned} \quad (12)$$

$$\begin{aligned}
 |\psi^+\rangle_{12}|\chi^+\rangle_{12} \rightarrow & \frac{i}{4}[(D_1'D_1' - A_1'A_1')(r_1'r_1' + l_1'l_1') \\
 & + (D_2'D_2' - A_2'A_2')(r_2'r_2' + l_2'l_2')]
 \end{aligned} \quad (13)$$

$$\begin{aligned}
 |\psi^-\rangle_{12}|\chi^+\rangle_{12} \rightarrow & \frac{i}{4}[(D_1'D_1' - A_1'A_1')(r_1'r_1' + l_1'l_1') \\
 & - (D_2'D_2' - A_2'A_2')(r_2'r_2' + l_2'l_2')]
 \end{aligned} \quad (14)$$

$$\begin{aligned}
 |\psi^+\rangle_{12}|\chi^-\rangle_{12} \rightarrow & \frac{i}{4}[(D_1'D_1' - A_1'A_1')(r_1'r_1' - l_1'l_1') \\
 & - (D_2'D_2' - A_2'A_2')(r_2'r_2' - l_2'l_2')]
 \end{aligned} \quad (15)$$

$$\begin{aligned}
 |\psi^-\rangle_{12}|\chi^-\rangle_{12} \rightarrow & \frac{i}{4}[(D_1'D_1' - A_1'A_1')(r_1'r_1' - l_1'l_1') \\
 & + (D_2'D_2' - A_2'A_2')(r_2'r_2' - l_2'l_2')]
 \end{aligned} \quad (16)$$

where $D = (|0\rangle^s + |1\rangle^s)/\sqrt{2}$, $A = (|0\rangle^s - |1\rangle^s)/\sqrt{2}$, $r = |0\rangle^o$ and $l = |1\rangle^o$. Then both of the output modes are passing through two polarizers at 45° . Only the output with the term $D_1'D_2'$ will result in there being one photon in each output mode with an efficiency of 1/8. The other terms will be rejected by the conditional detection. According to equations (1)–(16), 4 of the 16 output modes have the term $D_1'D_2'$ and the 4 corresponding input hyper-entangled Bell states are $|\phi^+\rangle_{12}|\omega^-\rangle_{12}$, $|\phi^-\rangle_{12}|\omega^+\rangle_{12}$, $|\phi^-\rangle_{12}|\chi^+\rangle_{12}$ and $|\phi^-\rangle_{12}|\chi^-\rangle_{12}$, which will be sent to the next stage of BSM on the OAM qubits.

BSM and teleportation of OAM qubits. Having measured and filtered out the SAM qubit, next we perform BSM on the OAM qubit. Dealing with a single degree of freedom is more straightforward, and can be implemented using a beam splitter. Like the PBS, the beam splitter also is not OAM-preserving, because the reflection at the beam splitter will flip the sign of OAM. Therefore, the transformation rules at a beam splitter for the four OAM Bell states are

$$|\omega^+\rangle_{12} \rightarrow \frac{i}{\sqrt{2}}(|0\rangle_1^o|1\rangle_1^o + |0\rangle_2^o|1\rangle_2^o)$$

$$|\omega^-\rangle_{12} \rightarrow \frac{1}{\sqrt{2}}(|0\rangle_1^o|0\rangle_2^o - |1\rangle_1^o|1\rangle_2^o)$$

$$\begin{aligned}
 |\chi^+\rangle_{12} \rightarrow & \frac{i}{2\sqrt{2}}(|0\rangle_1^o|0\rangle_1^o + |1\rangle_1^o|1\rangle_1^o \\
 & + |0\rangle_2^o|0\rangle_2^o + |1\rangle_2^o|1\rangle_2^o)
 \end{aligned}$$

$$\begin{aligned}
 |\chi^-\rangle_{12} \rightarrow & \frac{i}{2\sqrt{2}}(|0\rangle_1^o|0\rangle_1^o - |1\rangle_1^o|1\rangle_1^o \\
 & - |0\rangle_2^o|0\rangle_2^o + |1\rangle_2^o|1\rangle_2^o)
 \end{aligned}$$

We can see that only the Bell state $|\omega^-\rangle_{12}$ will result in there being one and only one photon in each output, whereas for the three other Bell states the two input photons will coalesce into a single output mode. Among these three Bell states, the state $|\omega^+\rangle_{12}$ can be further distinguished by measuring two single photons in

either output with the OAM orthogonal basis ($|0\rangle^o, |1\rangle^o$). Thus, this would allow us to unambiguously discriminate two from the four Bell states. Experimentally, we design dual-channel OAM readout devices to measure $|0\rangle^o$ and $|1\rangle^o$ simultaneously. Therefore, the efficiency of both QND and BSM on OAM is 1/2. The overall efficiency of h-BSM combining all three steps is $1/8 \times 1/2 \times 1/2 = 1/32$.

Generating three photon pairs. Ultrafast laser pulses with an average power of 800 mW, central wavelength of 394 nm, pulse duration of 120 fs and repetition rate of 76 MHz successively pass through three β -barium borate (BBO) crystals (Fig. 2) to generate three photon pairs through type-I spontaneous parametric down-conversion (SPDC). The first and the third crystals are 2 mm-thick BBOs, whereas the second one consists of two 0.6 mm-thick contiguous type-I BBO crystals with optic axes aligned in perpendicular planes. All of the down-converted photons have central wavelengths of 788 nm. The first photon pair (1–2) is initially prepared in the zero-order OAM mode with a coincidence count rate of 3.6×10^5 Hz. In SPDC, the zero-order OAM mode has the highest weight among all OAM modes and thus has the highest brightness. The second photon pair (2–3) is simultaneously entangled in both the SAM and the OAM (in the first-order mode), owing to OAM conservation in the type-I SPDC. It has a count rate of 3.4×10^4 Hz and a state fidelity of 0.95. The third photon pair (4–5) is created in the first-order OAM entangled state with a two-photon count rate of 1.2×10^5 Hz and a fidelity of 0.91. We estimate the mean numbers of photon pairs generated per pulse as $\sim 0.1, \sim 0.01$ and ~ 0.05 for the first, second and third pairs, respectively.

Preparing spin-orbit entanglement. As illustrated in the yellow panel of Fig. 2, we use a Sagnac interferometer to prepare the spin-orbit hybrid entangled state to be teleported, $|\phi\rangle_E = (|0\rangle^s|0\rangle^o + |1\rangle^s|1\rangle^o)/\sqrt{2}$. Photon 1 is initially prepared in the $(|0\rangle^s - |1\rangle^s)/\sqrt{2}$ SAM state with zero-order OAM. It is then sent into a Sagnac interferometer that consists of a PBS and an SPP, where the counter-propagating $|0\rangle^s$ and $|1\rangle^s$ SAMs with zero-order OAM pass through the SPP in opposite directions and are converted into $|0\rangle^s|0\rangle^o$ and $|1\rangle^s|1\rangle^o$, respectively. Considering an extra π phase shift for the $|1\rangle^s$ SAM from double reflections in the PBS, the final output state is $(|0\rangle^s|0\rangle^o + |1\rangle^s|1\rangle^o)/\sqrt{2}$. We note that the whole conversion process is deterministic.

Two-photon interference on the PBS and beam splitters. For a test of two-photon interference visibility at the PBS, input photons 1 and 2 are first intentionally prepared in the states $D_1r_1A_2l_2$ (orthogonal SAMs; see open squares in Extended Data Fig. 1a) and $D_1r_1D_2l_2$ (parallel SAMs; see solid circles in Extended Data Fig. 1a), respectively. We measure each single photon from the two outputs of the PBS in the D_1D_2 basis ($1'$ and $2'$ denote the output spatial modes). At zero delay, where the photons are optimally overlapped in time, the orthogonal SAM input yields an output state of $(|0\rangle_1^s|0\rangle_2^s + |1\rangle_1^s|1\rangle_2^s)/\sqrt{2}$ conditioned on a coincidence detection, which can be decomposed in the diagonal basis as $(|D\rangle_1^s|D\rangle_2^s + |A\rangle_1^s|A\rangle_2^s)/\sqrt{2}$, thus showing an enhancement. For the parallel SAM input, the output state is $(|D\rangle_1^s|A\rangle_2^s + |A\rangle_1^s|D\rangle_2^s)/\sqrt{2}$, which shows a reduction. The increase in the delay gradually destroys the indistinguishability of the two photons, such that their quantum state becomes a classical mixture. Thus, at large delays the counts appear flat. Interferometers of this type are sensitive only to length changes of the order of the coherence length of the detected photons and stay stable for weeks.

The two-photon interferences on beam splitters 1 and 2 are for the teleportation-based QND measurement of OAM qubits and the BSM of OAM qubits, respectively. As a test, the two input photons are prepared in the same SAM but orthogonal OAM states. It is interesting to note that, in stark contrast to the conventional Hong–Ou–Mandel interference, only having the two input OAM states orthogonal can lead to an interference dip, because the reflection at the beam splitter flips the sign of OAM. The two-photon interference as a function of temporal delay is shown in Extended Data Fig. 1b, c.

Dual-channel and efficient OAM measurement. One of the most frequently used OAM measurement devices in the previous experiments is off-axis hologram gratings^{21,23,24,31}, which typically have a practical efficiency (p) of about 30% (ref. 31). This low efficiency would cause an extremely low six-photon coincidence count rate that scales as p^6 , more challenging than the scaling (p^2) in the previous two-photon OAM experiments. To overcome this challenge, we use two different types of device for efficient OAM readout.

The first type is what we refer to as ‘dual-channel’ OAM measurement devices, used after the two beam splitters. The strategy is to transfer the OAM information to the photon’s SAM, and measure it using a PBS with two output channels. The method is as follows. After the beam splitter, each photon passes through a HWP and is prepared in the state $(|0\rangle^s + |1\rangle^s)/\sqrt{2}$. They are then sent into Sagnac interferometers with a Dove prism inside³² (Fig. 2), which is placed at a $\pi/8$ angle with respect to the interferometer plane. The photon is rotated by an angle of $\pi/4$ for the $|0\rangle^s$ SAM component when passing through the Dove prism forwards and by an angle of $-\pi/4$ for the $|1\rangle^s$ SAM component when passing through the Dove prism backwards. OAM states $|0\rangle^o$ and $|1\rangle^o$ with respective phases $e^{i\varphi}$ and $e^{-i\varphi}$, where φ is the azimuthal angle in the polar coordinate system, which pass through

a Dove prism rotated by $\pi/8$, will be rotated by an angle of $\pi/4$ and their phases will change to $e^{i(\varphi-\pi/4)}$ and $e^{-i(\varphi-\pi/4)}$, respectively. Therefore, two opposite phases, $e^{i\pi/4}$ and $e^{-i\pi/4}$, will be added to the two orthogonal OAM modes. Finally, the output photons are transformed into the $|0\rangle^s$ and $|1\rangle^s$ polarizations using a QWP. The overall transformations can be summarized as a CNOT gate between the OAM and SAM:

$$\begin{aligned} |0\rangle^s|0\rangle^o &\xrightarrow{\text{HWP}} (|0\rangle^s + |1\rangle^s)|0\rangle^o/\sqrt{2} \\ &\xrightarrow{\text{Sagnac}} e^{i\pi/4}(|0\rangle^s + i|1\rangle^s)|0\rangle^o/\sqrt{2} \\ &\xrightarrow{\text{QWP}} e^{i\pi/4}|0\rangle^s|0\rangle^o \\ |0\rangle^s|1\rangle^o &\xrightarrow{\text{HWP}} (|0\rangle^s + |1\rangle^s)|1\rangle^o/\sqrt{2} \\ &\xrightarrow{\text{Sagnac}} e^{-i\pi/4}(|0\rangle^s - i|1\rangle^s)|1\rangle^o/\sqrt{2} \\ &\xrightarrow{\text{QWP}} e^{-i\pi/4}|1\rangle^s|1\rangle^o \end{aligned}$$

Effectively, the OAM qubit is deterministically and redundantly encoded by the SAM qubit, that is, the SAM becomes identical to the OAM state. Thus, by measuring the SAM using a PBS with two output channels, we can recover information about the OAM with a high efficiency of $\sim 97\%$. In our experiment, four such Sagnac interferometers are used in the h-BSM. In this way, two of the four OAM Bell states can be discriminated using a beam splitter.

Whereas the first type is like a two-channel readout device (like a PBS for SAM), the second type is like a one-channel readout device (like a polarizer for SAM), and is used in the final stage of state verification after teleportation. The strategy for the projective OAM measurement is to transform it into the zero-order OAM mode so that the photon can be coupled into a single-mode fibre, while all other higher-order OAM modes will be rejected. Here we use an SPP³³ and a BPP³³ for efficient OAM readout.

The SPP³³ is designed with a spiral shape to create a vortex phase of $e^{il\varphi}$, where l is an integer and is referred to as topological charge. The SPP can attach vortex phases of $e^{il\varphi}$ and $e^{-il\varphi}$ to a photon when it passes through the SPP forwards and, respectively, backwards. When vortex phases of $e^{il\varphi}$ and $e^{-il\varphi}$ are attached, the corresponding OAM values will increase by and, respectively, decrease by lh . Therefore, a SPP can be used as an OAM mode converter. In our experiment, the OAM qubits are encoded in the OAM first-order subspace with the topological charge l being 1. The conversion between the OAM zero mode and the first-order modes ($|0\rangle^o$ and $|1\rangle^o$) is realized using a 16-phase-level SPP with an efficiency of $\sim 97\%$.

For the coherent transformation between the OAM zero mode and the superposition states $(|0\rangle^o \pm |1\rangle^o)/\sqrt{2}$ and $(|0\rangle^o \pm i|1\rangle^o)/\sqrt{2}$, a 2-phase-level BPP^{33,34} is used with an efficiency of $\sim 80\%$. These high-efficiency OAM measurement devices boost the sixfold coincidence count rate in our present experiment by more than two orders of magnitude, compared with the previous use of hologram gratings.

Error budget. The sources of error in our experiment include double pair emission in spontaneous parametric down-conversion; partial distinguishability of the independent photons that interfere at the PBS ($\sim 5\%$) and the beam splitters ($\sim 5\%$); state measurement error due to zero-order OAM leakage ($\sim 2\%$); fidelity imperfection of entangled photon pairs 2–3 ($\sim 5\%$) and 4–5 ($\sim 9\%$); and fidelity imperfection of the to-be-teleported single-photon hybrid entangled state ($\sim 8\%$).

Some error sources affect all the teleported states. First, the double pair emission contributed a $\sim 15\%$ background to the overall sixfold coincidence counts. If this were subtracted, the average teleportation fidelity would be improved to ~ 0.74 . Second, the imperfectly entangled photon pair 4–5 and the imperfect two-photon interference at beam splitters 1 and 2 (for the QND and OAM Bell-state measurements, respectively) degrade the teleportation fidelity for all states by $\sim 13\%$. Third, the imperfect state measurements mainly due to the zero-order OAM leakage cause a degradation of $\sim 2\%$.

Some error sources can have different effects on different teleported states. The imperfect two-photon interference at the PBS degrades the teleportation fidelities for the states $|\phi\rangle_C$, $|\phi\rangle_D$ and $|\phi\rangle_E$ by $\sim 5\%$. However, for the states $|\phi\rangle_A$ and $|\phi\rangle_B$ where the photon is horizontally or vertically polarized, the actual teleportation does not require two-photon interference at the PBS, and is therefore immune to the imperfection of the interference. This explains why the teleportation fidelities for the states $|\phi\rangle_A$ and $|\phi\rangle_B$ are the highest. This is inconsistent with the previous results: for the experiments using a PBS^{35–37}, the teleportation fidelities in the horizontal–vertical basis are higher than those in the $\pm 45^\circ$ linear and circular

polarization bases; whereas for the experiments^{38,39} using a non-polarizing beam splitter, the teleportation fidelities in all polarization bases are largely unbiased.

We note that the teleportation fidelity for the hybrid entangled state, $|\varphi\rangle_E$, is the lowest, which is affected by the imperfection ($\sim 8\%$) in the state preparation of the initial state $|\varphi\rangle_E$ as well as by imperfections in $|\varphi\rangle_C$ and $|\varphi\rangle_D$, that is, essentially all error and noise present in the experiment. It can be expected that all entangled states are subject to the same decoherence mechanism as the state $|\varphi\rangle_E$ as demonstrated, and should undergo similar reductions in fidelity.

These sources of noise can in principle be eliminated in future by various methods. For instance, deterministic entangled photons⁴⁰ do not suffer the problem of double pair emission. We also plan to develop bright OAM-entangled photons with higher fidelity, and a more precise 32-phase-level SPP for the next experiment of hyper-entanglement swapping.

A universal scheme for teleporting NDoFs. We illustrate in Extended Data Fig. 2 a universal scheme for teleporting N DoFs of a single photon. For simplicity, we discuss an example for three DoFs (Extended Data Fig. 2b), labelled X, Y and Z. There are in total 64 hyper-entangled Bell states:

$$\{|\phi\rangle_X^+, |\phi\rangle_X^-, |\psi\rangle_X^+, |\psi\rangle_X^-\} \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+, |\psi\rangle_Y^-\} \\ \otimes \{|\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+, |\psi\rangle_Z^-\}$$

These the products of the Bell states of each DoF, defined as

$$|\phi\rangle_i^\pm = (|0\rangle_i |0\rangle_i \pm |1\rangle_i |1\rangle_i) / \sqrt{2}$$

$$|\psi\rangle_i^\pm = (|0\rangle_i |1\rangle_i \pm |1\rangle_i |0\rangle_i) / \sqrt{2}$$

where $i = X, Y, Z$. The aim is to perform an h-BSM, identifying 1 out of the 64 states. The required resources include photon pairs entangled in the Z DoF, pairs hyper-entangled in the Y-Z DoFs, pairs hyper-entangled in the X-Y-Z DoFs, filters for the three DoFs that can project the state to $|0\rangle$ or $|1\rangle$ (with a functionality similar to the polarizers for SAM), qubit flip (σ_x) operations for the DoFs, a 50:50 non-polarizing beam splitter and single-photon detectors, all of which are commercially available or have been experimentally demonstrated previously.

It has been known that if two single photons are superposed at a beam splitter, only asymmetric quantum states can result in one and only one photon exiting from each output of the splitter. In the simplest case of one DoF, this is the asymmetric $|\psi\rangle^-$ state. As now we have three DoFs, we have to consider the combined, molecular-like quantum states. There are in total 28 possible combinations that are asymmetric states, as follows:

$$|\psi\rangle_X^- \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\} \otimes \{|\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+\} \\ \{|\phi\rangle_X^+, |\phi\rangle_X^-, |\psi\rangle_X^+\} \otimes |\psi\rangle_Y^- \otimes \{|\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+\} \\ \{|\phi\rangle_X^+, |\phi\rangle_X^-, |\psi\rangle_X^+\} \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\} \otimes |\psi\rangle_Z^- \\ |\psi\rangle_X^- \otimes |\psi\rangle_Y^- \otimes |\psi\rangle_Z^-$$

After the photons have passed through the first beam splitter, we can filter out and retain these 28 from the 64 hyper-entangled Bell states conditioned on seeing one and only one photon in each output of the splitter. Next we apply two filters in the two output of the splitter to project the X DoF into $|\psi\rangle^\pm$. One filter is set to pass the $|1\rangle$ state and the other is set to pass the $|0\rangle$ state. This results in the following 16 states being filtered from the 28:

$$|\psi\rangle_X^- \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\} \otimes \{|\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+\} \\ |\psi\rangle_X^+ \otimes |\psi\rangle_Y^- \otimes \{|\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+\} \\ |\psi\rangle_X^+ \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\} \otimes |\psi\rangle_Z^- \\ |\psi\rangle_X^- \otimes |\psi\rangle_Y^- \otimes |\psi\rangle_Z^-$$

We perform a bit-flip operation on the X DoF on one of the arms of the interferometer, erasing the information in the X DoF. We then pass them into the second beam splitter, filtering out and retaining the six asymmetric combinations

$$|\psi\rangle_X^+ \otimes \{|\psi\rangle_Y^-, |\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+\}$$

$$|\psi\rangle_X^+ \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\} \otimes |\psi\rangle_Z^-$$

We emphasize that, before sending the photons into the second beam splitter, for the reason discussed in the main text we use teleportation-based QND measurement to ensure the two photons can be fed into the subsequent cascaded interferometers. Here the QND should preserve the quantum information in the Y and Z DoFs. Thus, quantum teleportation of two DoFs of a single photon is required (Extended Data Fig. 2a), which is exactly what we demonstrated in the experiment presented in the main text.

After that, we again pass the photons through two filters on the Y DoF, one set in the $|1\rangle$ state and the other set in the $|0\rangle$ state, which leaves four asymmetric combinations:

$$|\psi\rangle_X^+ \otimes \{|\psi\rangle_Y^-, |\phi\rangle_Z^+, |\phi\rangle_Z^-, |\psi\rangle_Z^+\} \\ |\psi\rangle_X^+ \otimes |\psi\rangle_Y^+ \otimes |\psi\rangle_Z^-$$

Similarly, we perform a bit-flip operation on the Y DoF on one of the arms, erasing the information in the Y DoF. We then do a QND measurement on the Z DoF, and pass the two photons into the third beam splitter, finally filtering out the only remaining asymmetric state:

$$|\psi\rangle_X^+ \otimes |\psi\rangle_Y^+ \otimes |\psi\rangle_Z^-$$

By detecting one and only one photon in the output of the third splitter, we can thus discriminate this particular hyper-entangled state, $|\psi\rangle_X^+ \otimes |\psi\rangle_Y^+ \otimes |\psi\rangle_Z^-$, from the 64 hyper-entangled Bell states on three DoFs.

To experimentally demonstrate the teleportation of three DoFs, the scheme would need in total ten photons (or five entangled photon pairs from SPDC), which is within the reach of near-future experimental abilities, given the recent advances in high-efficiency photon collection and detection. It is obvious that the above protocol can be extended to more DoFs as displayed in Extended Data Fig. 2c.

Feed-forward scheme for spin-orbit composite states. To realize a deterministic teleportation, feed-forward Pauli operations on the teleported particle based on the intrinsically random Bell-state measurement results are essential. In our present experiment, no feed-forward has been applied. Here we briefly describe how this can be done for the spin-orbit composite state. For the SAM qubits, active feed-forward has been demonstrated before using fast electro-optical modulators^{39,41} (EOMs). To take advantage of this technology, which has been demonstrated to have high-speed operation and high gate fidelity, we use a coherent quantum SWAP gate between the OAM and SAM qubits. The SWAP gate is defined as

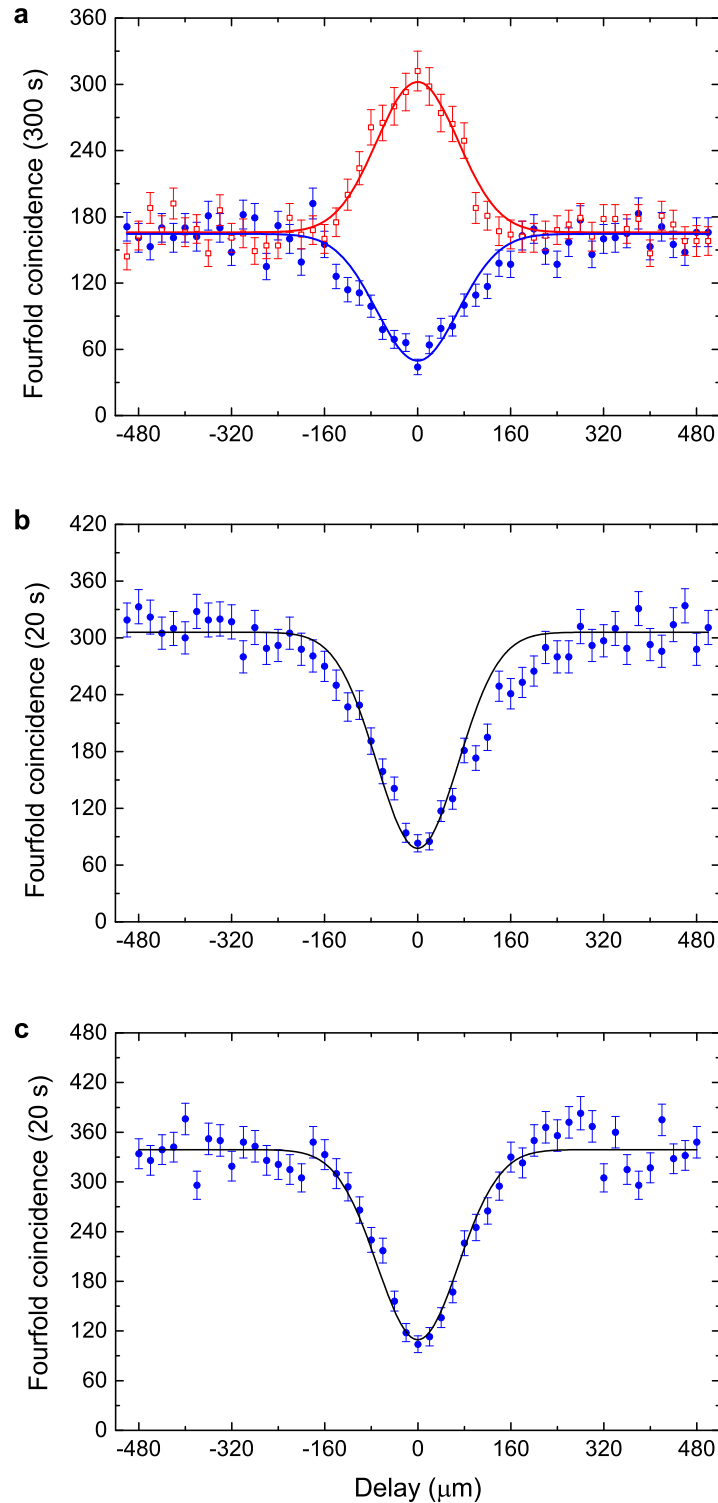
$$\alpha|0\rangle^s|0\rangle^o + \beta|1\rangle^s|0\rangle^o + \gamma|0\rangle^s|1\rangle^o + \delta|1\rangle^s|1\rangle^o \\ \xrightarrow{\text{SWAP}} \alpha|0\rangle^o|0\rangle^s + \beta|1\rangle^o|0\rangle^s + \gamma|0\rangle^o|1\rangle^s + \delta|1\rangle^o|1\rangle^s$$

The operation sequence for the feed-forward operation on the spin-orbit composite states is shown in Extended Data Fig. 3a. First, the SAM feed-forward is done with an EOM. Second, the SAM and OAM qubits undergo a SWAP operation. Third, an EOM is used to operate on the 'new' SAM that is converted from the OAM. We note that the EOM does not affect the OAM, and so the previous operation on the SAM is unaffected. Lastly, a final SWAP gate converts the SAM back to OAM, which completes the feed-forward for spin-orbit composite states.

The SWAP gate is composed of three CNOT gates (Extended Data Fig. 3b). In the first and third CNOT gates (blue shading), the SAM qubit is the control qubit that acts on the OAM as target qubit, realized by sending a photon through, and recombined at a PBS. On the reflection arm only, a Dove prism is inserted to induce an OAM bit flip. In the second CNOT (yellow shading), the control and target qubits are respectively the SAM and OAM qubits, as explained in the Methods section 'Dual-channel and efficient OAM measurement'. The extra Dove prism and HWP in front of the PBS are used to compensate for the phase shift inside the Sagnac interferometer.

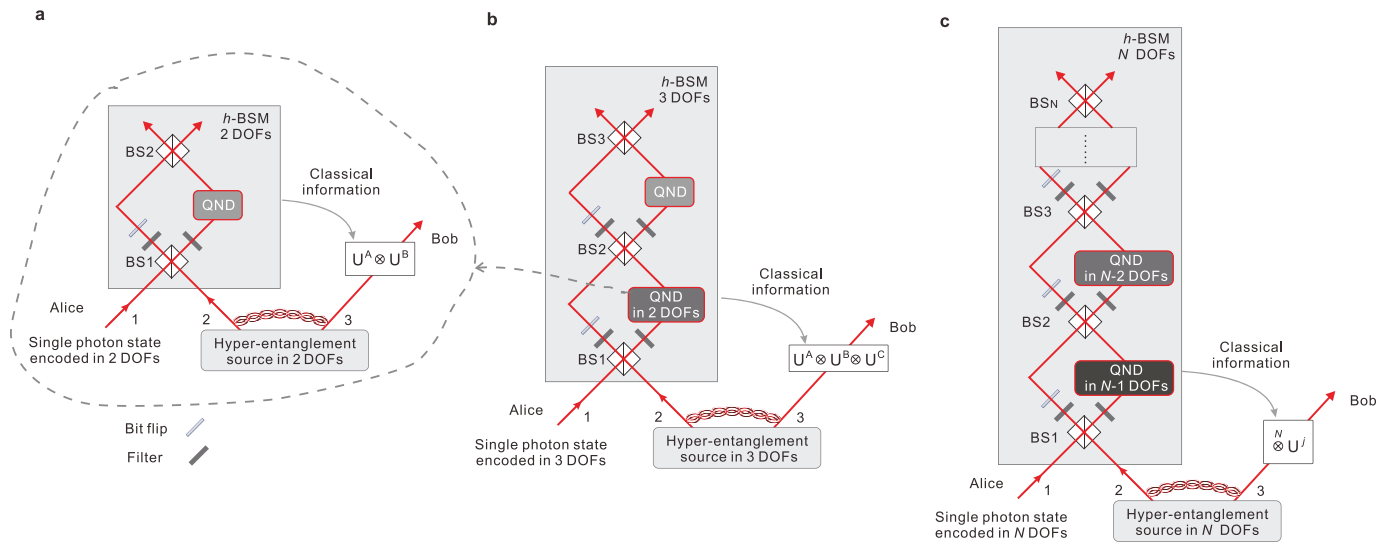
- Graham, T. M., Barreiro, J. T., Mohseni, M. & Kwiat, P. G. Hyperentanglement-enabled direct characterization of quantum dynamics. *Phys. Rev. Lett.* **110**, 060404 (2013).
- Slussarenko, S. *et al.* The polarizing Sagnac interferometer: a tool for light orbital angular momentum sorting and spin-orbit photon processing. *Opt. Express* **18**, 27205–27216 (2010).
- Andrews, D. L. & Babiker, M. (eds) *The Angular Momentum of Light* (Cambridge Univ. Press, 2012).

34. Jack, B. *et al.* Precise quantum tomography of photon pairs with entangled orbital angular momentum. *New J. Phys.* **11**, 103024 (2009).
35. Pan, J.-W., Gasparoni, S., Aspelmeyer, M., Jennewein, T. & Zeilinger, A. Experimental realization of freely propagating teleported qubits. *Nature* **421**, 721–725 (2003).
36. Zhang, Q. *et al.* Experimental quantum teleportation of a two-qubit composite system. *Nature Phys.* **2**, 678–682 (2006).
37. Yin, J. *et al.* Quantum teleportation and entanglement distribution over 100-kilometre free-space channels. *Nature* **488**, 185–188 (2012).
38. Bouwmeester, D. *et al.* Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
39. Ma, X.-S. *et al.* Quantum teleportation over 143 kilometres using active feed-forward. *Nature* **489**, 269–273 (2012).
40. Lu, C.-Y. & Pan, J.-W. Push-button photon entanglement. *Nature Photon.* **8**, 174–176 (2014).
41. Prevedel, R. *et al.* High-speed linear optics quantum computing using active feed-forward. *Nature* **445**, 65–69 (2007).



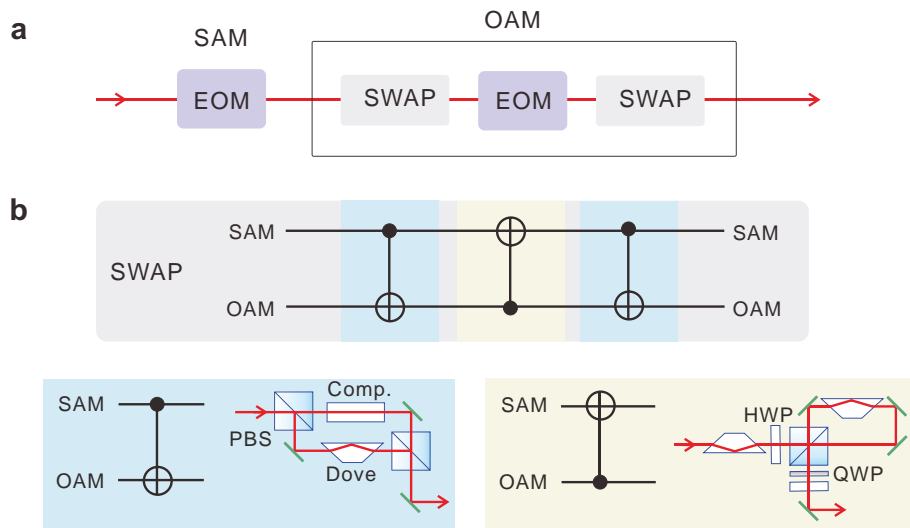
Extended Data Figure 1 | Hong–Ou–Mandel interference of multiple independent photons encoded with SAM or OAM. **a**, Interference at the PBS where input photons 1 and 2 are intentionally prepared in the states $D_1r_1A_2l_2$ (orthogonal SAMs; open squares) and $D_1r_1D_2l_2$ (parallel SAMs; solid circles). The y axis shows the raw fourfold (the trigger photon t and photons 1, 2 and 3) coincidence counts. The extracted visibility is 0.75 ± 0.03 , calculated from $V(0) = (C_+ - C_{//})/(C_+ + C_{//})$, where $C_{//}$ and C_+ are the coincidence counts without any background subtraction at zero delay for parallel and, respectively, orthogonal SAMs. The red and blue lines are Gaussian fits to the raw data. **b**, Two-photon interference on beam splitter 1, where photons 1 and 4 are

prepared in orthogonal OAM states. The black line is a Gaussian fit to the raw data of fourfold (the trigger photon and photons 1, 4 and 5) coincidence counts. The visibility is 0.73 ± 0.03 , calculated from $V(0) = 1 - C_0/C_\infty$, where C_0 and C_∞ is the fitted counts at zero and, respectively, infinite delays. **c**, Two-photon interference at beam splitter 2, where input photons 1 and 5 are prepared in the orthogonal OAM states. The black line is a Gaussian fit to the data points. The interference visibility is 0.69 ± 0.03 calculated in the same way as in **b**. Error bars, 1 s.d., calculated from Poissonian counting statistics of the raw detection events.



Extended Data Figure 2 | A universal scheme for teleporting N DoFs of a single photons. **a**, A scheme for teleporting two DoFs of a single photon using three beam splitters, which is slightly different from the one presented in the main text using a PBS and two beam splitters. Through the first beam splitter, six asymmetric states, $|\psi\rangle_X^- \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\}$ and $\{|\phi\rangle_X^+, |\phi\rangle_X^-, |\psi\rangle_X^+\} \otimes |\psi\rangle_Y^-$, can result in one photon in each output, which is ensured by teleportation-based QND on the Y DoF. After passing the two photons through the two filters that project them into the $|1\rangle$ and $|0\rangle$ states for the X DoF, four states, $|\psi\rangle_X^- \otimes \{|\phi\rangle_Y^+, |\phi\rangle_Y^-, |\psi\rangle_Y^+\}$ and $|\psi\rangle_X^+ \otimes |\psi\rangle_Y^-$, survive. Through the second beam splitter, only the asymmetric state $|\psi\rangle_Y^-$ of the Y DoF can result in one photon in each output. Finally we can discriminate the state

$|\psi\rangle_X^+ \otimes |\psi\rangle_Y^-$ from the 16 hyper-entangled Bell states. **b**, Teleportation of three DoFs of a single photons (Methods). Note that to ensure that there is one and only one photon in the output of the first beam splitter, we can use the teleportation-based QND on two DoFs in **a** (dashed circle). **c**, Generalized teleportation of N DoFs of a single photons. The h-BSM on N DoFs can be implemented as follows: (1) the beam splitter post-selects the asymmetric hyper-entangled Bell states in N DoFs which contain an odd number of asymmetric Bell states in one DoF, (2) two filters and one bit-flip operation erase the information on the measured DoF and further post-select asymmetric states, and (3) teleportation-based QND.



Extended Data Figure 3 | Active feed-forward for spin-orbit composite states. **a**, The active feed-forward scheme. This composite active feed-forward could be completed in a step-by-step manner. First, we use an EOM to implement the active feed-forward for SAM qubits. It is important to note that EOM does not affect OAM. Second, we use a coherent quantum SWAP gate between the OAM and SAM qubits. The original OAM is converted into a 'new' SAM, whose active feed-forward operation is done by a second EOM. Then the

OAM and SAM qubits undergo a second SWAP operation and are converted to the original DoFs. **b**, The quantum circuit for a SWAP gate between the OAM and SAM qubits. The SWAP gate is composed of three CNOT gates: in the first and third CNOT gates, the SAM and OAM qubits act as the control and target qubits, respectively, whereas in the second CNOT gate this is reversed.

Dynamically reconfigurable complex emulsions via tunable interfacial tensions

Lauren D. Zarzar¹, Vishnu Sresht², Ellen M. Sletten¹, Julia A. Kalow¹, Daniel Blankschtein² & Timothy M. Swager¹

Emulsification is a powerful, well-known technique for mixing and dispersing immiscible components within a continuous liquid phase. Consequently, emulsions are central components of medicine, food and performance materials. Complex emulsions, including Janus droplets (that is, droplets with faces of differing chemistries) and multiple emulsions, are of increasing importance¹ in pharmaceuticals and medical diagnostics², in the fabrication of microparticles and capsules^{3–5} for food⁶, in chemical separations⁷, in cosmetics⁸, and in dynamic optics⁹. Because complex emulsion properties and functions are related to the droplet geometry and composition, the development of rapid, simple fabrication approaches allowing precise control over the droplets' physical and chemical characteristics is critical. Significant advances in the fabrication of complex emulsions have been made using a number of procedures, ranging from large-scale, less precise techniques that give compositional heterogeneity using high-shear mixers and membranes¹⁰, to small-volume but more precise microfluidic methods^{11,12}. However, such approaches have yet to create droplet morphologies that can be controllably altered after emulsification. Reconfigurable complex liquids potentially have great utility as dynamically tunable materials. Here we describe an approach to the one-step fabrication of three- and four-phase complex emulsions with highly controllable and reconfigurable morphologies. The fabrication makes use of the temperature-sensitive miscibility of hydrocarbon, silicone and fluorocarbon liquids, and is applied to both the microfluidic and the scalable batch production of complex droplets. We demonstrate that droplet geometries can be alternated between encapsulated and Janus configurations by varying the interfacial tensions using hydrocarbon and fluorinated surfactants. This yields a generalizable strategy for the fabrication of multiphase emulsions with controllably reconfigurable morphologies and the potential to create a wide range of responsive materials.

Phase separation approaches using mass transfer of a co-solvent or separating agent^{13–16} have attracted interest as simplified routes to the fabrication of complex emulsions. In designing our new method, we use the facts that fluorocarbons are lipophobic as well as hydrophobic and that many fluorocarbon and hydrocarbon liquids are immiscible at room temperature but have a low upper consolute temperature (T_c) and mix with gentle heating¹⁷. Hexane and perfluorohexane, for example, have a T_c of 22.65 °C (ref. 18). Our interest in fluorocarbons and their emulsions was also based on the fact that they are inert materials with unique properties that have been exploited for use as magnetic resonance imaging and ultrasound contrast agents, as artificial blood, in water-repellent surfaces, and in the acoustically triggered release of payloads^{17,19,20}. To explore the feasibility of using a temperature-induced phase separation route to complex emulsions (Fig. 1a), we emulsified a 1:1 volume ratio of hexane and perfluorohexane above T_c in an aqueous solution of Zonyl FS-300 (hereafter 'Zonyl'), which is a nonionic fluorosurfactant with the linear chemical formula $F(CF_2)_xCH_2CH_2O(CH_2CH_2O)_yH$. Cooling below T_c induced phase separation and yielded structured complex droplets (Fig. 1b). These complex emulsions were readily produced

in bulk by shaking warm hexane–perfluorohexane liquid in a surfactant solution (Fig. 1c). Although these droplets were polydisperse, the morphology and composition of the droplets were highly uniform. Chemical partitioning during phase separation¹⁴ gave directed compartmentalization of solutes (Fig. 1d). Therefore, temperature-induced phase separation of liquids provides a simple, scalable approach to the fabrication of complex functional emulsions.

The morphology of the complex droplets is exclusively controlled by interfacial tension. To put this in an analytical context, consider a complex emulsion of any immiscible liquids F and H (at a given volume ratio) in a third immiscible liquid W. We choose the interfacial tensions of the H–W interface, γ_{HW} , the F–W interface, γ_{FW} , and the F–H interface, γ_{FH} , such that γ_{FW} and γ_{HW} are significantly larger than γ_{FH} . This regime is relevant to combinations of liquids H and F that have low interfacial tension just below T_c . It can be shown²¹ that these multiphase droplets are nearly spherical in shape and will adopt one of the following three thermodynamically permissible internal configurations: (1) liquid H completely encapsulates liquid F, (2) liquids H and F form a Janus droplet, and (3) liquid F completely encapsulates liquid H (Fig. 2a). These droplet configurations are characterized by two contact angles, θ_{HW} between the H–W and F–H interfaces, and θ_{FW} between the F–W and F–H interfaces. The three interfacial tensions acting along the

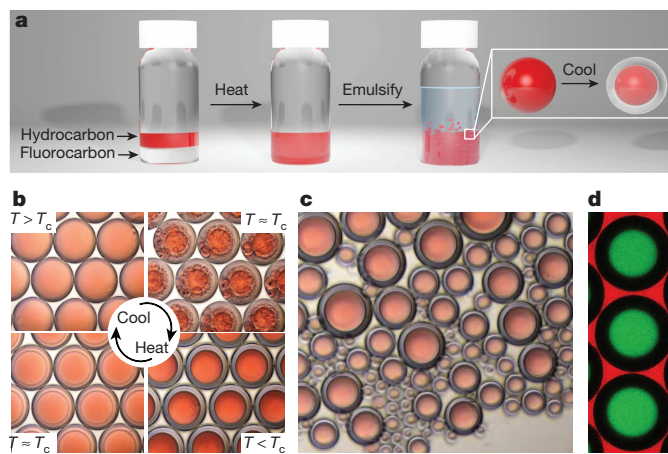


Figure 1 | Temperature-controlled phase separation of hydrocarbon and fluorocarbon liquids can be used to create complex emulsions. **a**, Complex emulsion fabrication. **b**, Above T_c , hexane and perfluorohexane are miscible and emulsified in aqueous 0.1% Zonyl (top left). Below T_c , hexane and perfluorohexane phases separate to create a hexane-in-perfluorohexane-in-water (H/F/W) double emulsion (bottom right). Hexane is dyed red. Scale bar, 200 μ m. **c**, Emulsions of uniform composition made by bulk emulsification (such as shaking). Scale bar, 100 μ m. **d**, Lateral confocal cross-section of H/F/W double-emulsion droplets. Hydrocarbon-soluble Nile Red dye (green) selectively extracts into hexane. Rhodamine B dyes the aqueous phase (red). Scale bar, 100 μ m. Monodisperse droplets in **b** and **d** were made using a microcapillary device.

¹Department of Chemistry and Institute for Soldier Nanotechnologies, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ²Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

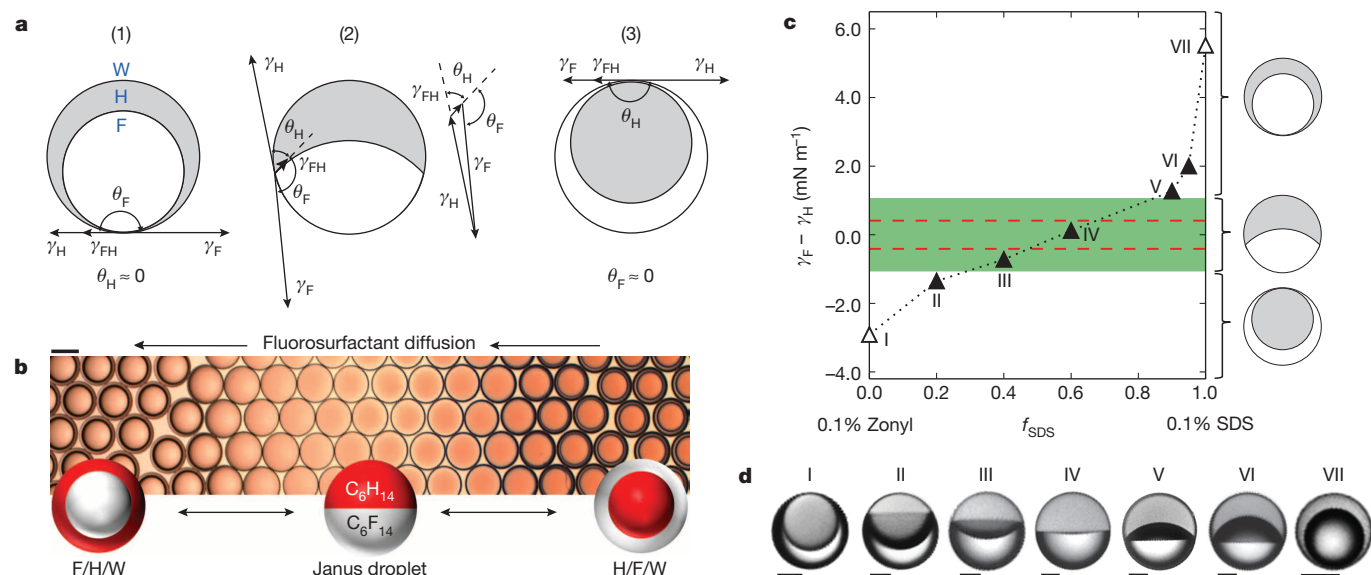


Figure 2 | Reconfiguration of droplet morphology is dynamic and results from changes in the balance between interfacial tensions. **a**, Sketch of the effect of interfacial tensions on the configuration of a complex droplet. In (1), $\gamma_F > \gamma_H + \gamma_{\text{FH}}$, favouring encapsulation of phase F within phase H. In (3), $\gamma_H > \gamma_F + \gamma_{\text{FH}}$ and phase H is encapsulated within phase F. At intermediate values of γ_F and γ_H , a Janus droplet with geometry typified by (2) is formed. γ_F , γ_H and γ_{FH} can be reconfigured into a Neumann triangle²⁹ solvable for θ_H and θ_F . **b**, Hexane-perfluorohexane droplets reconfigure in response to variation in the concentration of Zonyl as it diffuses through 0.1% SDS from right (higher Zonyl concentration) to left (lower Zonyl concentration). Scale bar, 100 μm . **c**, Configurational stability diagram for the

hexane-perfluorohexane-water system showing $\gamma_F - \gamma_H$ as a function of the fraction of 0.1% SDS, f_{SDS} , where the other fraction is 0.1% Zonyl. The green band denotes the region $|\gamma_F - \gamma_H| < \gamma_{\text{FH}} = 1.07 \pm 0.1$ mN m⁻¹, obtained from geometrical analysis of Janus droplets. The red dashed lines correspond to $\gamma_{\text{FH}} = 0.4$ mN m⁻¹ as predicted for a temperature of 10 °C (ref. 23), the approximate temperature at which the droplets in **d** were imaged. Filled and unfilled triangles indicate conditions under which Janus droplets and double emulsions were observed, respectively. Labels I–VII correspond to the droplets in **d**. **d**, Optical micrographs of hexane-perfluorohexane droplets in solutions of 0.1% Zonyl and 0.1% SDS in varying ratios as plotted in **c**. Hexane is dyed and appears grey. Scale bars, 50 μm .

interfaces must be in equilibrium for the droplet configuration to be stable, as can be expressed by the following equations:

$$\cos(\theta_H) = \frac{\gamma_F^2 - \gamma_H^2 - \gamma_{\text{FH}}^2}{2\gamma_{\text{FH}}\gamma_H} \quad (1)$$

$$\cos(\theta_F) = \frac{\gamma_H^2 - \gamma_F^2 - \gamma_{\text{FH}}^2}{2\gamma_{\text{FH}}\gamma_F} \quad (2)$$

Figure 2a shows that configurations (1) and (3) are the limiting cases of configuration (2) as $\theta_H \rightarrow 0$ and $\theta_F \rightarrow 0$, respectively. Equations (1) and (2) can be used to translate these limiting contact angles into interfacial tension conditions, yielding the following two relationships:

$$\theta_H = 0 \Rightarrow \gamma_F = \gamma_H + \gamma_{\text{FH}} \quad (3)$$

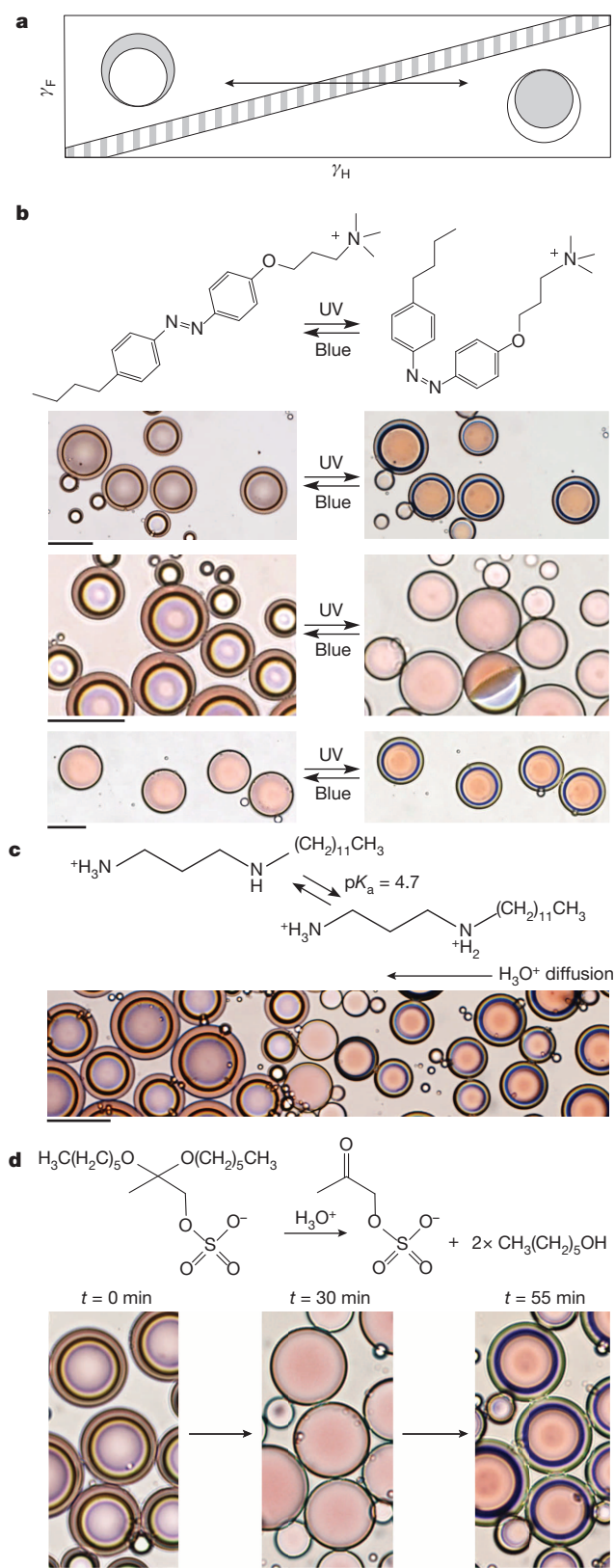
$$\theta_F = 0 \Rightarrow \gamma_H = \gamma_F + \gamma_{\text{FH}} \quad (4)$$

Recast as the difference between γ_H and γ_F , equations (3) and (4) indicate that when $\gamma_F - \gamma_H \geq \gamma_{\text{FH}}$ the droplets assume configuration (1) in Fig. 2a. Conversely, when $\gamma_H - \gamma_F \geq \gamma_{\text{FH}}$ the droplets adopt configuration (3) in Fig. 2a. However, when the difference between γ_H and γ_F is of the order of γ_{FH} , the droplets adopt a Janus droplet geometry associated with configuration (2) in Fig. 2a.

These physical relationships reveal that, given a low value of γ_{FH} , only slight changes in the balance of γ_H and γ_F are necessary to induce dramatic changes in the droplet's morphology. We proposed that if liquids F, H and W were fluorocarbon, hydrocarbon and water, respectively, dynamic reconfiguration of the droplets might be triggered by hydrocarbon and fluorinated surfactants²². Consistently, a 1:1 volume mixture of hexane and perfluorohexane in 0.1% Zonyl generated hexane-in-perfluorohexane-in-water (H/F/W) double emulsions (Fig. 1), indicating a preferential decrease in γ_F . In comparison, emulsification of the same hexane-perfluorohexane mixture in water containing 0.1% sodium dodecyl sulphate (SDS, an anionic hydrocarbon surfactant), yielded

F/H/W double emulsions as a result of a preferentially decreased γ_H . When we introduced a small volume of 10% Zonyl into F/H/W droplets in 0.1% SDS, we observed that the droplet morphology dynamically changed in accordance with the concentration gradient of Zonyl. Droplets first passed through a spherical Janus drop morphology before inverting to H/F/W double emulsions (Fig. 2b and Supplementary Video 1). The reverse was seen when concentrated SDS was added to 0.1% Zonyl-stabilized droplets (Supplementary Video 2). These results suggest that not only is droplet morphology highly controllable on initial emulsification, by the choice of surfactants, but that these emulsions are also dynamically and reversibly reconfigurable by changing the balance of γ_F and γ_H .

To validate the proposed dynamic mechanism, we measured the tensions of hexane-water (γ_H) and perfluorohexane-water (γ_F) interfaces for a variety of 0.1% SDS and 0.1% Zonyl ratios using the pendant-drop method (Methods, Extended Data Fig. 1 and Extended Data Table 1). The perfluorohexane-hexane interfacial tension, γ_{FH} , has been estimated previously²³ and is in general agreement with our γ_{FH} estimates based on geometrical analysis of the Janus drops produced (Methods, Extended Data Fig. 2 and Extended Data Table 2). We can use the quantity $\gamma_F - \gamma_H$ as a simple indicator of droplet configuration, and this quantity is plotted as a function of surfactant ratio in Fig. 2c. When only 0.1% SDS was used ($f_{\text{SDS}} = 1$), we measured $\gamma_F - \gamma_H > \gamma_{\text{FH}}$ and observed complete encapsulation of perfluorohexane by hexane. In a mixed SDS-Zonyl composition, the trajectory entered the narrow green zone (Fig. 2c) corresponding to $|\gamma_F - \gamma_H| < \gamma_{\text{FH}}$. Within this Janus droplet configurational zone, the droplet morphology began to 'flip' as visualized in Fig. 2d. As the proportion of 0.1% SDS approached zero and fluoro-surfactant dominated, the trajectory left the Janus zone and the droplets assumed a configuration in which hexane was encapsulated by perfluorohexane. Overall, the observed droplet geometries closely followed the predicted morphological trend and confirm that variations in hydrocarbon and fluorinated surfactants are effective for the manipulation of the balance between γ_F and γ_H .



This understanding has made it possible for us to induce droplet morphological transitions with stimuli-responsive and cleavable surfactants (Fig. 3a). Responsive surfactants²⁴ undergo reversible changes in their effectiveness²⁵ when triggered by stimuli such as magnetic fields, pH, CO_2 or light. To create optically sensitive droplets, we synthesized a light-responsive surfactant²⁶ consisting of an azobenzene moiety that

Figure 3 | Emulsions reconfigure in response to light and pH. **a**, Sketch of how the variations in γ_F and γ_H induced by alterations in the effectiveness of a hydrocarbon surfactant translate into differences in drop morphology on a phase-stability diagram. Grey represents hexane and white represents perfluorohexane. **b**, Chemical structure of the light-responsive surfactant which reversibly isomerizes under ultraviolet (UV) and blue light between the more effective *trans* form of the surfactant (left) and the less effective *cis* form (right). Aligned beneath are optical micrographs of hexane-perfluorohexane emulsions that are tuned to undergo specific morphological transitions in response to light. Hexane is dyed red, and the aqueous phase consists of Zonyl and the light-responsive surfactant pictured. Top: droplets undergo complete inversion. Middle: F/H/W double-emulsion drops transition to Janus droplets. Most droplets are viewed from the top, but one is lying on its side allowing a view of the droplet profile. Bottom: Janus droplets transition to an H/F/W double emulsion. Scale bar, 100 μm . **c**, A pH-responsive surfactant, *N*-dodecylpropane-1,3-diamine, is used in combination with Zonyl to create pH-responsive droplets. Acid diffuses through the solution from right (higher concentration) to left (lower concentration), reducing the pH below $\text{pK}_a = 4.7$ and thereby generating a less effective surfactant³⁰ and inducing inversion of the emulsion from F/H/W to H/F/W. Scale bar, 100 μm . **d**, Emulsions stabilized by a combination of Zonyl and acid-cleavable surfactant, sodium 2,2-bis(hexyloxy)propyl sulphate²⁸, undergo morphological changes as the cleavable surfactant is degraded over time at pH = 3. Scale bar, 50 μm .

reversibly undergoes a photo-induced isomerization between a more effective *trans* configuration and a less effective *cis* configuration (Fig. 3b). When this surfactant was used in combination with Zonyl, we observed that the hexane-perfluorohexane droplets rapidly and reversibly changed morphology in response to ultraviolet (wavelength, $\lambda = 365$ nm) and blue ($\lambda = 470 \pm 20$ nm) light. Depending on the relative concentrations of Zonyl and the light-responsive surfactant and on the duration or intensity of light exposure, we tuned the morphology to switch between the double-emulsion and Janus states or to invert entirely (Supplementary Videos 3 and 4). Analogous results were achieved with Zonyl and a pH-responsive surfactant, *N*-dodecylpropane-1,3-diamine, by alternating the pH between values above and below the surfactant's lowest pK_a value, 4.7 (ref. 27; Fig. 3c). Similarly, cleavable surfactants show a reduction in efficacy when irreversibly degraded by exposure to light, heat or changes in pH. In this context, we demonstrated irreversible transition between the F/H/W and H/F/W double-emulsion states (Fig. 3d) using an acid-cleavable surfactant²⁸, sodium 2,2-bis(hexyloxy)propyl sulphate, in conjunction with Zonyl. The results shown here demonstrate the versatility of the complex fluorocarbon-hydrocarbon droplets' response to a wide range of stimuli.

Liquid droplets and solid particles with asymmetric properties were created by effecting different chemistries in the separate compartments of a fluorocarbon-hydrocarbon, or more generally fluorous-organic, Janus droplet. To create directionally orientable and movable liquid Janus droplets, we synthesized magnetic Fe_3O_4 nanoparticles stabilized with oleic acid for preferential partitioning into the organic phase. Janus droplets with hemispheres of ethyl nonafluorobutyl ether (the fluorous phase) and dichlorobenzene with Fe_3O_4 (the organic phase) were rapidly oriented and moved in the direction of a magnet (Fig. 4a). To generate solid hemispherical particles, we polymerized an emulsion consisting of a liquid polymer precursor, 1,6-hexanediol diacrylate, as the organic phase and methoxyperfluorobutane as the fluorous phase (Fig. 4b). By replacing methoxyperfluorobutane with a fluorinated acrylate oligomer and crosslinker, we created spherical solid Janus particles with fluorinated and non-fluorinated sides (Fig. 4c).

The same principles of droplet transformations observed in three-phase emulsions were extended to a four-phase system, thereby generating reconfigurable droplets of even higher-order complexity. We designed a system comprised of silicone oil (Si), hydrocarbon oil (H, mineral oil and octadecane) and fluorinated oil (F, ethyl nonafluorobutyl ether) such that the liquids mixed when heated and separated into three phases at room temperature (20 $^\circ\text{C}$). Systematically varying ratios of 1% Zonyl and 1% SDS aqueous surfactant solutions caused droplets to assume morphologies combining both Janus configurations (indicated

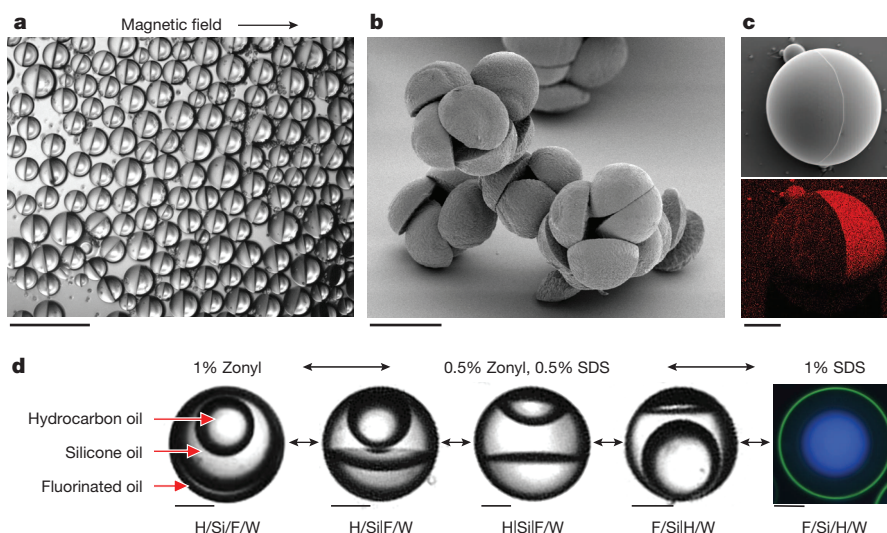


Figure 4 | Magnetic complex emulsions, complex emulsions as templates and four-phase emulsions. **a**, Janus droplets of ethyl nonafluorobutyl ether and dichlorobenzene containing magnetite nanoparticles in the dichlorobenzene phase are oriented with a magnet. Scale bar, 200 μm . **b**, Scanning electron micrograph of hemispherical particles made from photopolymerized Janus droplets containing hexanediol diacrylate and methoxyperfluorobutane. Scale bar, 100 μm . **c**, Top: scanning electron micrograph of a Janus particle with hydrocarbon and fluorinated polymeric hemispheres. Bottom: the energy-dispersive X-ray spectral map reveals the fluorinated hemisphere. Scale bar, 50 μm . **d**, Four-phase emulsions reconfigure

in response to changes in interfacial tension. Drops contain hydrocarbon oil (H, mineral oil with octadecane), silicone oil (Si) and fluorinated oil (F, ethyl nonafluorobutyl ether) emulsified in water (W). The silicone phase is enriched with a fraction of the two other phases. Left: 1% Zonyl generates an H/Si/F/W triple emulsion. Right: 1% SDS generates an F/Si/H/W triple emulsion with a thin outer shell. Phases were identified with fluorescent dyes. In the fluorescence micrograph, Nile Red (green) preferentially dyes the hydrocarbon oil, and fluoros-tagged coumarin dye³¹ preferentially dyes the fluoros phase (blue). Scale bars, 50 μm .

using ‘|’) and encapsulated configurations (indicated using ‘/’) (Fig. 4d). For example, a triple emulsion H/Si/F/W was formed in 1% Zonyl, but in a 3:2 ratio of 1% Zonyl:1% SDS we observed a Janus configuration between the fluoros and silicone phases while the hydrocarbon phase remained encapsulated in the silicone phase, H/Si|F/W. As the proportion of SDS increased, the droplet passed through Janus droplet and mixed Janus–encapsulated droplet configurations before inverting to the reverse triple emulsion (F/Si/H/W) in 1% SDS.

Complex droplets of controllable composition and dynamic reconfigurable morphology provide a new active element for novel and existing applications of emulsions. The fabrication method and the dynamic mechanism presented are general and can be broadly applied using a wide variety of chemicals, materials and surfactants well beyond the initial demonstrations described here. Droplets triggered by environmental stimuli could be used, for example, to target the release of drugs at tumours, to induce changes in colour or transparency for camouflage, as vehicles for the sequestration of pollutants, as tunable lenses, or as sensors. Emulsions with the characteristic ability to selectively ‘present’ and ‘hide’ specific liquid interfaces and controllably alter droplet morphology and symmetry will find abundant applications.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 October; accepted 24 December 2014.

- Aserin, A. *Multiple Emulsions: Technology and Applications* (Wiley, 2008).
- Gresham, P. A., Barnett, M. & Smith, S. V. Use of a sustained-release multiple emulsion to extend the period of radioprotection conferred by cysteamine. *Nature* **234**, 149–150 (1971).
- Lone, S. & Cheong, I. W. Fabrication of polymeric Janus particles by droplet microfluidics. *RSC Adv.* **4**, 13322–13333 (2014).
- Kim, J.-W., Utada, A. S., Fernández-Nieves, A., Hu, Z. & Weitz, D. A. Fabrication of monodisperse gel shells and functional microgels in microfluidic devices. *Angew. Chem. Int. Ed.* **119**, 1851–1854 (2007).
- Shum, H. C. *et al.* Droplet microfluidics for fabrication of non-spherical particles. *Macromol. Rapid Commun.* **31**, 108–118 (2010).
- Augustin, M. A. & Hemar, Y. Nano- and micro-structured assemblies for encapsulation of food ingredients. *Chem. Soc. Rev.* **38**, 902–912 (2009).
- Chakravarti, A., Chowdhury, S., Chakrabarty, S., Chakrabarty, T. & Mukherjee, D. Liquid membrane multiple emulsion process of chromium (VI) separation from waste waters. *Colloids Surf. A* **103**, 59–71 (1995).
- Patravale, V. & Mandawgade, S. Novel cosmetic delivery systems: an application update. *Int. J. Cosmet. Sci.* **30**, 19–33 (2008).
- Chen, L. *et al.* Photoresponsive monodisperse cholesteric liquid crystalline microshells for tunable omnidirectional lasing enabled by a visible light-driven chiral molecular switch. *Adv. Optical Mater.* **2**, 845–848 (2014).
- van der Graaf, S., Schroën, C. G. P. H. & Boom, R. M. Preparation of double emulsions by membrane emulsification—a review. *J. Membr. Sci.* **251**, 7–15 (2005).
- Shah, R. K. *et al.* Designer emulsions using microfluidics. *Mater. Today* **11**, 18–27 (2008).
- Choi, C.-H. *et al.* Microfluidic design of complex emulsions. *ChemPhysChem* **15**, 21–29 (2014).
- Zhao, C.-X. & Middelberg, A. P. J. Microfluidic mass-transfer control for the simple formation of complex multiple emulsions. *Angew. Chem. Int. Ed.* **48**, 7208–7211 (2009).
- Choi, C. H., Weitz, D. A. & Lee, C. S. One step formation of controllable complex emulsions: from functional particles to simultaneous encapsulation of hydrophilic and hydrophobic agents into desired position. *Adv. Mater.* **25**, 2536–2541 (2013).
- Haase, M. F. & Bruij, J. Tailoring of high-order multiple emulsions by the liquid–liquid phase separation of ternary mixtures. *Angew. Chem. Int. Ed.* **53**, 11793–11797 (2014).
- Song, Y. & Shum, H. C. Monodisperse w/w/w double emulsion induced by phase separation. *Langmuir* **28**, 12054–12059 (2012).
- Gładysz, J. A., Curran, D. P. & Horváth, I. T. *Handbook of Fluorous Chemistry* 18–20, 521–561 (Wiley-VCH, 2004).
- Bedford, R. G. & Dunlap, R. D. Solubilities and volume changes attending mixing for the system: perfluoro-*n*-hexane-*n*-hexane. *J. Am. Chem. Soc.* **80**, 282–285 (1958).
- Wong, T.-S. *et al.* Bioinspired self-repairing slippery surfaces with pressure-stable omniphobicity. *Nature* **477**, 443–447 (2011).
- Duncanson, W. J. *et al.* Microfluidic fabrication of perfluorohexane-shelled double emulsions for controlled loading and acoustic-triggered release of hydrophilic agents. *Langmuir* **30**, 13765–13770 (2014).
- Guzowski, J., Korczyk, P. M., Jakiela, S. & Garstecki, P. The structure and stability of multiple micro-droplets. *Soft Matter* **8**, 7269–7278 (2012).
- Mukerjee, P. & Handa, T. Adsorption of fluorocarbon and hydrocarbon surfactants to air-water, hexane-water and perfluorohexane-water interfaces. Relative affinities and fluorocarbon-hydrocarbon nonideality effects. *J. Phys. Chem.* **85**, 2298–2303 (1981).
- McClain, B., Yoon, M., Litster, J. & Mochrie, S. Interfacial roughness in a near-critical binary fluid mixture: X-ray reflectivity and near-specular diffuse scattering. *Eur. Phys. J. B* **10**, 45–52 (1999).
- Brown, P., Butts, C. P. & Eastoe, J. Stimuli-responsive surfactants. *Soft Matter* **9**, 2365–2374 (2013).
- Rosen, M. J. & Kunjappu, J. T. *Surfactants and Interfacial Phenomena* 238 (Wiley, 2012).

26. Chevallier, E. *et al.* Pumping-out photo-surfactants from an air–water interface using light. *Soft Matter* **7**, 7866–7874 (2011).
27. Li, X., Yang, Y., Eastoe, J. & Dong, J. Rich self-assembly behavior from a simple amphiphile. *ChemPhysChem* **11**, 3074–3077 (2010).
28. Li, M., Powell, M. J., Razunguzwa, T. T. & O'Doherty, G. A. A general approach to anionic acid-labile surfactants with tunable properties. *J. Org. Chem.* **75**, 6149–6153 (2010).
29. Rowlinson, J. S. & Widom, B. *Molecular Theory of Capillarity* 209–212 (Clarendon, 1982).
30. Yang, Y. *et al.* Environmentally responsive adsorption and assembly behaviors from *N*-alkyl-1,2-ethylenediamines. *Soft Matter* **9**, 1458–1467 (2013).
31. Sletten, E. M. & Swager, T. M. Fluorofluorophores: fluorescent fluororous chemical tools spanning the visible spectrum. *J. Am. Chem. Soc.* **136**, 13574–13577 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements Financial support from Eni S.p.A. under the Eni-MIT Alliance Solar Frontiers Program and by the US Army Research Laboratory and the US Army

Research Office through the Institute for Soldier Nanotechnologies under contract number W911NF-13-D-0001 is acknowledged. E.M.S. and J.A.K. were supported by F32 Ruth L. Kirschstein NRSA Fellowships under award numbers EB014682 (E.M.S.) and GM106550 (J.A.K.). We thank L. Arriaga for introducing us to the fabrication of capillary microfluidic devices.

Author Contributions L.D.Z. and T.M.S. developed the concept for the research. L.D.Z. conducted experiments involving emulsion fabrication and imaging, and measured interfacial tensions. V.S. and D.B. modelled the system and calculated and analysed equilibrium interfacial tensions. E.M.S. synthesized the fluorinated crosslinker and fluorinated coumarin dye. J.A.K. synthesized the light-responsive surfactant and the cleavable surfactant. All authors contributed to the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.B. (dblank@mit.edu) or T.M.S. (tswager@mit.edu).

METHODS

Chemicals. The following chemicals were used as received: sodium dodecyl sulphate ($\geq 99\%$), Sudan Red 7B (95%), 1,6-hexanediol diacrylate (80%), 1,4-butanediol diacrylate (90%), trimethylolpropane ethoxylate triacrylate ($M_n = 428 \text{ g mol}^{-1}$), 1,2-dichlorobenzene (99%), Zonyl FS-300 (40% solids), methoxyperfluorobutane (99%), mineral oil (light), iron(II) chloride tetrahydrate (99.99%), 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9-hexadecafluorodecane-1,10-diol (97%), methacrylic acid (99%) and octadecane (99%) (Sigma-Aldrich); silicone oil (for oil baths -40 to $+200^\circ\text{C}$), perfluorohexanes (98%), hexanes (98%), and iron(III) chloride (98%) (Alfa Aesar); ethyl nonafluorobutyl ether ($>98\%$) and sodium oleate ($>97\%$) (TCI); Darocur 1173 (Ciba); fluorinated acrylate oligomer (Sartomer); Nile Red (99%) (Acros); *N*-dodecylpropane-1,3-diamine ($>95\%$) (Matrix). Light-responsive surfactant was made by the literature procedure^{26,32}. Fluorinated coumarin dye was made by the literature procedure³¹. Cleavable surfactant sodium 2,2-bis(hexyloxy)propyl sulphate was made by the literature procedure²⁸.

Synthesis of the fluorinated crosslinker. To synthesize the fluorinated crosslinker 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9-hexadecafluorodecane-1,10-diyl bis(2-methylacrylate), a mixture of 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9-hexadecafluorodecane-1,10-diol (10 g, 22 mmol, 1 equiv.), methacrylic acid (20 ml, 240 mmol, 11 equiv.), 2,6-di-*tert*-butyl-4-methylphenol (200 mg, 0.9 mmol, 0.04 equiv.) and sulphuric acid (0.1 ml, 18 M) in toluene (250 ml) was heated to reflux with azeotropic removal of water (Dean–Stark trap). After three days, the mixture was cooled to room temperature and washed with saturated aqueous sodium bicarbonate ($3 \times 300 \text{ ml}$). The toluene was dried with MgSO_4 and evaporated. The remaining residue/product was dissolved in perfluorohexanes (150 ml) and filtered. The filtrate was evaporated to yield 7.81 g of colourless liquid (13.1 mmol, 61% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3): δ 6.22 (p, $J = 1.1 \text{ Hz}$, 2H), 5.70 (p, $J = 1.5 \text{ Hz}$, 2H), 4.65 (t, $J = 13.3 \text{ Hz}$, 4H), 1.98 (t, $J = 1.3 \text{ Hz}$, 6H). $^{13}\text{C-NMR}$ (101 MHz, CDCl_3): δ 165.8, 135.0, 127.9, 117.5–107.5 (m, CF_2), 60.1 (t, $J = 27.3 \text{ Hz}$), 18.1. $^{19}\text{F-NMR}$ (376 MHz, CDCl_3): δ -119.3 (p, $J = 13.3 \text{ Hz}$, 4F), -121.8–-122.0 (m, 8F), -123.3 (bs, 4F). LRMS (EI): calculated for $\text{C}_{18}\text{H}_{14}\text{F}_{16}\text{O}_4 [\text{M}]^+$, 598; found, 598. NMR spectra were obtained on a Bruker Avance 400 MHz spectrometer. LRMS was acquired on an Agilent 5973N GCMS. Please see Extended Data Fig. 3 and Extended Data Fig. 4 for reaction scheme and NMR spectra.

General fabrication of complex emulsions. The hydrocarbon and fluorocarbon liquids of choice were heated until miscible and emulsified. The temperature required varied depending on the solutions. Solutions were emulsified either in bulk by shaking or by coaxial glass capillary microfluidics and cooled to induce phase separation. For hexane–perfluorohexane emulsions, the emulsions were chilled on ice before imaging and often imaged while immersed in a cool water bath to maintain a temperature below 20°C . For microfluidics, Harvard Apparatus PHD Ultra syringe pumps were used to inject the outer phase and inner phase using a glass capillary microfluidic device made from an outer square capillary (outer diameter, 1.5 mm; inner diameter, 1.05 mm; AIT Glass) and inner cylindrical capillary (outer diameter, 1 mm; World Precision Instruments) pulled to a $30 \mu\text{m}$ tip using a P-1000 Micropipette Puller (Sutter Instrument Company). The microfluidic set-up was heated above the T_c of the inner phase solution using a heat lamp. Emulsions were then cooled below T_c to induce phase separation. Emulsions were observed to be stable during the time periods used (of the order of days). Longer-term stability experiments were not conducted.

Measurement of interfacial tensions. Interfacial tension measurements were made using the pendant-drop method (ramé-hart Model 500 Advanced Goniometer). Measurements on a drop were taken every 5 s until the interfacial tension appeared to be nearing equilibrium or the droplet became unstable. The hexane–water interfacial tension was measured to be 50 mN m^{-1} and the perfluorohexane–water interfacial tension was measured to be 55 mN m^{-1} .

Microscopy. Lateral confocal cross-sections of the droplets were imaged using a Nikon 1AR ultrafast spectral scanning confocal microscope. Scanning electron microscopy was conducted on gold-sputtered samples with a JEOL 6010LA scanning electron microscope. Fluorescence and bright-field images were taken with a Zeiss Axiovert 200 inverted microscope equipped with a Zeiss AxioCam HRc camera. Droplets typically orient themselves with the denser, fluorine phase downward. To take side-view images of the drops, emulsions were shaken to induce the drops to roll around while images were made with a 1 ms exposure.

Fabrication of magnetic Janus droplets. Magnetite nanoparticles were made as follows: 25 ml of concentrated NH_3OH was added to an acidified solution of 1.6 g of FeCl_3 and 1 g of $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$ in 50 ml of water at 80°C . The magnetite nanoparticle precipitate was collected with a magnet, washed with water and redispersed. One gram of sodium oleate in 10 ml of water was added while stirring at room temperature. The oily black precipitate was extracted with hexanes. The solid was collected by evaporation of solvent, and was subsequently redispersed in dichlorobenzene. Janus droplets were obtained by heating the nanoparticle–dichlorobenzene solution

and ethyl nonafluorobutyl ether above T_c and shaking in 0.2% SDS and 0.2% Zonyl in a 2.5:1 ratio. The drops were oriented using a neodymium magnet.

Fabrication of light-responsive emulsions. Hexane dyed with Sudan Red 7b and perfluorohexane in equal volumes were used as the inner phase, and a mixture of 0.1% light-sensitive surfactant and 0.1% Zonyl FS-300 in an 8:2 ratio was used as surfactants in the aqueous phase. Slight adjustments to the surfactant concentrations were made during imaging to tune the droplet morphological transitions to achieve the desired outcome; for example, more Zonyl was needed to generate droplets that transitioned from a Janus droplet to a hexane–perfluorohexane–water double emulsion. A mercury lamp was used as the intense light source, and DAPI and FITC filters were used to selectively allow ultraviolet ($\lambda = 365 \text{ nm}$) and blue ($\lambda = 470 \pm 20 \text{ nm}$) light to reach the sample while imaged on an inverted microscope.

Fabrication of reversibly responsive pH-sensitive emulsions. Hexane and perfluorohexane in equal volumes were emulsified in a solution of 2 mM *N*-dodecylpropane-1,3-diamine and 1.2% Zonyl in 0.2 M NaCl. Salt solution was used to control for effects of ionic strength on the emulsion morphology. The pH was adjusted to above and below pH 4.7 by addition of HCl and NaOH.

Fabrication of emulsions using an acid-cleavable surfactant. Hexane and perfluorohexane in equal volumes were emulsified in a solution of 0.3% sodium 2,2-bis(hexyloxy)propyl sulphate and 0.4% Zonyl in 0.1 M NaCl. The pH was adjusted to 3 using HCl and the solution was allowed to sit undisturbed on the inverted microscope stage while images were periodically made over an hour.

Fabrication of hemispherical particles. 1,6-hexanediol diacrylate with 4% Darocur 1173 photo-initiator was heated with an equal volume of methoxyperfluorobutane above T_c and emulsified. 1% SDS and 1% Zonyl in a 3:2 ratio yielded Janus droplets which were then polymerized under a Dymax Blue Wave 200 ultraviolet lamp while kept cold on ice.

Fabrication of fluorine–hydrocarbon Janus particles. Fluorinated acrylate oligomer, fluorinated crosslinker, 1,4-butanediol diacrylate and trimethylolpropane ethoxylate triacrylate were used in a volume ratio of 15:3:14:10. 5% Darocur 1173 was used as the photo-initiator. The inner phase mixture was heated above T_c , emulsified in 1% SDS and polymerized with ultraviolet light over ice.

Fabrication of four-phase emulsions. Light mineral oil with 20 wt% octadecane (used to reduce T_c in a mixture with the other liquids), silicone oil and ethyl nonafluorobutyl ether were used as the inner phases in a volume ratio of 6:7:13. The mineral oil and ethyl nonafluorobutyl ether both partition into the silicone oil such that on phase separation the silicone oil phase is enriched with some quantity of the two other phases. Aqueous mixtures of varying ratios of 1% Zonyl and 1% SDS were used as the outer phase, and emulsions were formed in bulk by shaking.

Estimation of equilibrium interfacial tensions in the presence of surfactants. Experimentally, it proved difficult to measure the interfacial tensions with the pendant-drop method accurately for very long periods of time while maintaining a stable drop volume because surfactant continued to adsorb at the droplet interface, further reducing the already low interfacial tensions. Therefore, it is likely that the final interfacial tensions, γ , recorded for the hexane–water interface and, respectively, the perfluorohexane–water interface for different surfactant compositions, f_{SDS} , as illustrated in Extended Data Fig. 1a, are not the equilibrium interfacial tensions, γ_{eqb} , for that system. In the absence of additional experimental data for long time scales, we used theories of dynamic interfacial tension to estimate γ_{eqb} .

Rosen suggested³³ that the dynamic interfacial tension of surfactants could be accurately modelled using an empirical model of the form

$$\gamma(t) = \gamma_{\text{eqb}} + \frac{\gamma_0 - \gamma_{\text{eqb}}}{1 + (t/t^*)^n} \quad (5)$$

where $\gamma_0 = \gamma(t = 0)$ and t^* and n are positive, empirically determined constants. The value of n depends on the type of surfactant and the interface considered, as well as on the concentration of the surfactant, and must therefore be independently determined for every set of experimental conditions. The constant t^* is the half-life of the interfacial tension decay process—it is the time taken for the difference, $\Delta\gamma = \gamma(t) - \gamma_{\text{eqb}}$, to decrease to half its initial value. The three unknown model parameters (γ_{eqb} , t^* and n) can be determined from a least-squares fit of equation (5) to the experimental dynamic interfacial tension data. The measure of the ‘goodness-of-fit’ of equation (5) to the experimental data is quantified by the coefficient of determination, or R^2 , value³⁴. The closer this value is to 1, the better the fit.

An example of the fit of the model to the dynamic interfacial tension data is shown in red in Extended Data Fig. 1a. The parameter results of using equation (5) to fit the interfacial tension data for all the surfactant systems considered are presented in Extended Data Table 1. The smallest value of R^2 obtained for the experimental data across all the systems studied was 0.909, indicating very good agreement between the predictions made using equation (5) and the experimental data generated using the goniometer.

The resulting extracted values of γ_{eqb} have been used to plot the variation in the interfacial tensions γ_{H} and γ_{F} as a function of the 0.1% SDS fraction, f_{SDS} , in Fig. 2

and Extended Data Fig. 1b. We note that some droplets fall just outside the designated Janus region. Discrepancies may be due to our inability to measure and estimate the equilibrium interfacial tensions with the 0.1 mN m^{-1} accuracy required, or there may be slight mixing of hexane and perfluorohexane in the droplets, which would alter the interfacial tensions relative to those of the pure liquids.

Estimation of γ_{FH} from the analysis of Janus droplet images. Bulk mixtures of hexane and perfluorohexane have an upper consolute temperature, T_c , of approximately 22°C , with a critical density of 1.14 g cm^{-3} and a volume ratio very close to 1:1 (ref. 18). The interfacial tension of the hexane–perfluorohexane interface at temperatures close to (but below) this upper consolute temperature is close to zero, and, consequently, it is difficult to measure accurately using conventional laboratory techniques such as pendant-drop or Du Nuoy ring tensiometry. The best estimate in the physical chemistry literature comes from the predictions of a model for capillary fluctuations at an interface fitted to X-ray reflectivity measurements²³. In ref. 23 it is proposed that the variation in the interfacial tension γ_{FH} (in units of mN m^{-1}) with temperature T (in units of K) is given by

$$\gamma(T) = 23.2 \left(\frac{T_c - T}{T_c} \right)^{1.26} \quad (6)$$

Equation (6) predicts that at a temperature of 283.15 K (10°C), the hexane–perfluorohexane interfacial tension should be 0.4 mN m^{-1} . We were unable to find other experimental data in the literature to support this estimate. However, given the equilibrium interfacial tension values and images of the resulting Janus droplets, it was possible to also estimate γ_{FH} for our system.

The relationships between the shapes of Janus droplets and the ratio of the volumes of the two constituent phases, k , and the relative magnitudes of the interfacial tensions operating at the various interfaces of the droplet have been examined in ref. 21. The theoretical treatment there can be used to relate the radii of curvature of the F–H, H–W and F–W interfaces to the volume ratio and the interfacial tensions γ_{H} , γ_{F} and γ_{FH} :

$$\frac{\gamma_{\text{H}}}{R_{\text{H}}} - \frac{\gamma_{\text{F}}}{R_{\text{F}}} = \frac{\gamma_{\text{FH}}}{R_{\text{FH}}} \quad (7)$$

Given the interfacial tension values at the H–W and F–W interfaces (γ_{H} and γ_{F} , respectively), equation (7) can be used to determine γ_{FH} if we can independently calculate the three radii of curvature R_{H} , R_{F} and R_{FH} .

The shape of a typical hexane–perfluorohexane Janus droplet in water is illustrated in Extended Data Fig. 2a. The droplet's interfaces are spherical arcs with different radii of curvature (denoted by R_{H} , R_{F} and R_{FH}). Furthermore, the line of contact between the two phases on the outer (water-facing) surface of the droplet is a circle whose diameter, D , is indicated by the dotted line in Extended Data Fig. 2a, b. The droplet in Extended Data Fig. 2a can be deconstructed into three spherical caps of the type shown in Extended Data Fig. 2b, each with the same base—a circle of diameter D —and respective radii R_{H} , R_{F} and R_{FH} . The volumes of these three spherical caps are denoted by $V_{\text{cap}}(R_{\text{H}}, D)$, $V_{\text{cap}}(R_{\text{F}}, D)$ and $V_{\text{cap}}(R_{\text{FH}}, D)$, respectively.

From the definition of the volume ratio, k , between the hexane and perfluorohexane phases, it follows that:

$$k = \frac{V_{\text{F}}}{V_{\text{H}}} = \frac{V_{\text{cap}}(R_{\text{F}}, D) - V_{\text{cap}}(R_{\text{FH}}, D)}{V_{\text{cap}}(R_{\text{H}}, D) + V_{\text{cap}}(R_{\text{FH}}, D)} \quad (8)$$

The volume of a spherical cap, $V_{\text{cap}}(r, d)$, can be calculated using the principles of elementary three-dimensional geometry³⁵. Knowledge of the values of k , R_{H} , R_{F} and

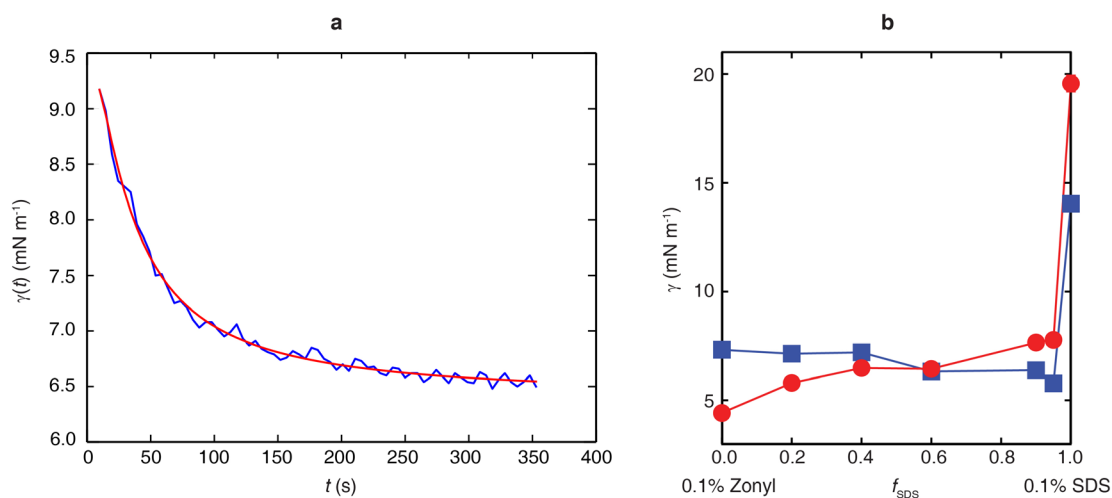
D allows us to solve equation (8) for the radius of curvature of the F–H interface, R_{FH} . Equation (7) can then be used to solve for the missing interfacial tension, γ_{FH} .

The two radii of curvature, R_{H} and R_{F} , and the diameter, D , of the circle of contact can be computed by analysing the images of the Janus droplets. A schematic of the analysis process is shown in Extended Data Fig. 2c–e. A raw image of a Janus droplet (Extended Data Fig. 2c) is first subjected to the Canny edge-detection algorithm³⁶ based on the non-normalized contrast, with a Gaussian kernel radius of 2.0 and low and high thresholds of 2.0 and 7.0, respectively. This procedure results in the detection of the outer (water-facing) interfaces of the Janus droplet as shown in Extended Data Fig. 2d. Arcs corresponding to the H–W and F–W interfaces are then selected by visual inspection and fitted to circles using Taubin's algorithm³⁷. These arcs and their fitted circles are depicted as blue and red circles (with green dots for centres) in Extended Data Fig. 2e. The radii of these two circles correspond to the radii of curvature, R_{H} and R_{F} , of the H–W and F–W interfaces, respectively. The diameter, D , of the circle of contact between the hexane and perfluorohexane phases is calculated as the minimum separation between the points lying between the arcs chosen to represent the H–W and F–W interfaces. This line of minimum separation is shown in Extended Data Fig. 2e. Image preprocessing and edge detection were performed in ImageJ v1.48, and the circle-fitting and interfacial tension calculation was performed in MATLAB.

This droplet analysis was performed for six Janus droplets, three from each of the two conditions $f_{\text{SDS}} = 0.4$ and $f_{\text{SDS}} = 0.6$. The interfacial tensions obtained are tabulated in Extended Data Table 2. The values of γ_{FH} do not vary significantly between the two sets of surfactant systems, which is consistent with our expectation that γ_{FH} is at best only weakly affected by the composition of the surfactant. The average value of γ_{FH} deduced from this analysis is $1.07 \pm 0.1 \text{ mN m}^{-1}$.

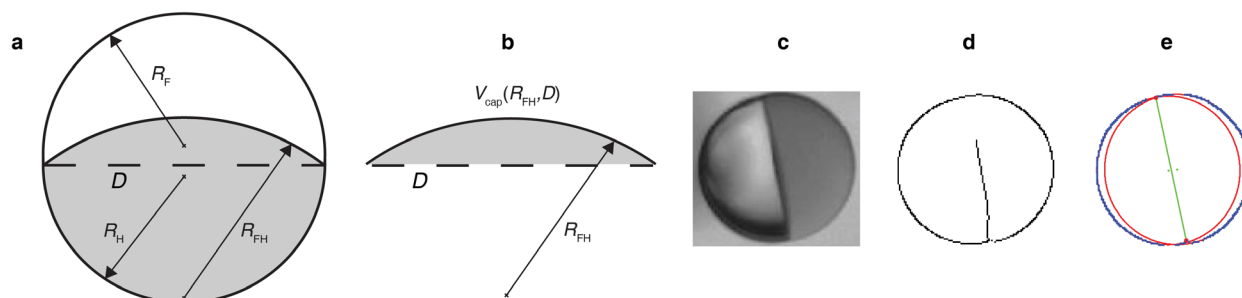
We note that accurately determining the interfacial tensions through the radii of curvature of the interfaces of the droplet is feasible only in the absence of fluid flow around the complex emulsion droplets. The presence of such a flow, even at low shear rates, leads to a deformation of the surface area of the droplet, causing local gradients in surfactant concentration and inducing Marangoni stresses³⁸. The interplay between these effects ultimately determines the interfacial tensions at any point on a complex emulsion droplet. Higher shear rates can dramatically disrupt droplet stability, leading to the fragmentation of the complex emulsion and the release of the encapsulated phases—a phenomenon whose onset has been examined in detail by theoretical³⁹ and experimental⁴⁰ studies.

32. Haiying, L. & Zhongfan, L. A convenient synthesis of novel mercapto-ended azobenzene derivatives. *Synth. Commun.* **28**, 3779–3785 (1998).
33. Hua, X. Y. & Rosen, M. J. Dynamic surface tension of aqueous surfactant solutions: 1. Basic parameters. *J. Colloid Interface Sci.* **124**, 652–659 (1988).
34. Glantz, S. A. & Slinker, B. K. *Primer of Applied Regression and Analysis of Variance* 2nd edn, 248 (McGraw-Hill, 1990).
35. Harris, J. W. & Stöcker, H. *Handbook of Mathematics and Computational Science* 107 (Springer, 1998).
36. Canny, J. A computational approach to edge detection. *IEEE Trans. Patt. Anal. Mach. Interf.* **PAMI-8**, 679–698 (1986).
37. Taubin, G. Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Trans. Patt. Anal. Mach. Interf.* **13**, 1115–1138 (1991).
38. Fischer, P. & Erni, P. Emulsion drops in external flow fields — the role of liquid interfaces. *Curr. Opin. Colloid Interface Sci.* **12**, 196–205 (2007).
39. Stone, H. A. & Leal, L. G. Breakup of concentric double emulsion droplets in linear flows. *J. Fluid Mech.* **211**, 123–156 (1990).
40. Muguet, V. et al. W/O/W multiple emulsions submitted to a linear shear flow: correlation between fragmentation and release. *J. Colloid Interface Sci.* **218**, 335–337 (1999).



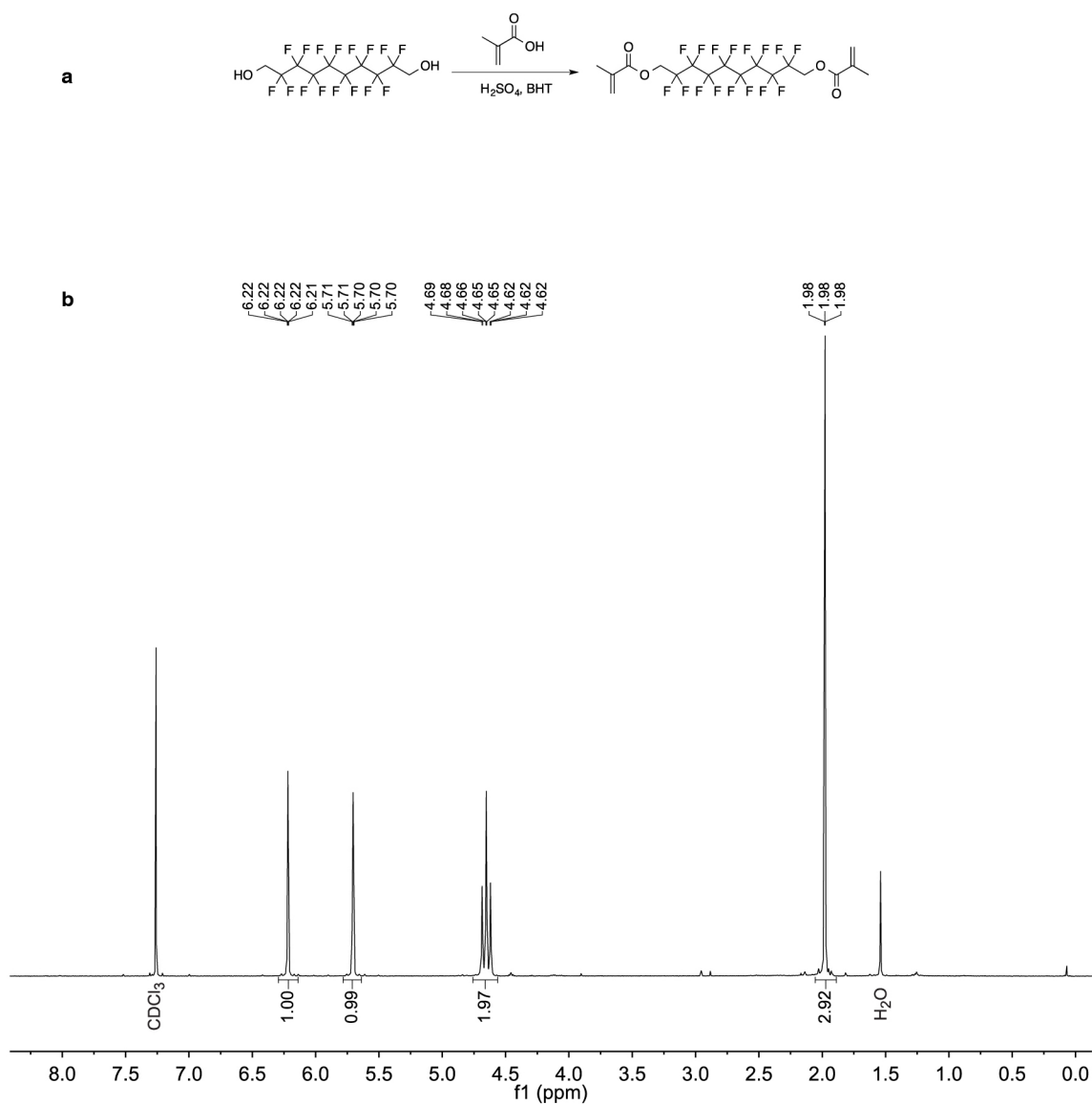
Extended Data Figure 1 | Dynamic interfacial tension data was used to estimate the equilibrium interfacial tensions for the hexane–water and perfluorohexane–water interfaces. **a**, Dynamic interfacial tension data (in blue) was obtained from the pendant-drop method; the representative data shown here was measured for the hexane–water interface at $f_{\text{SDS}} = 0.9$ (such that the aqueous solution contained 0.1% SDS and 0.1% Zonyl in a 9:1 ratio). The data was fitted to an empirical model (in red) to estimate the

equilibrium value of the interfacial tension $\gamma_{\text{eqb}} = \gamma(t \rightarrow \infty)$. Such fitting was performed for all measured interfacial tensions and the fitted parameter results are tabulated in Extended Data Table 1. **b**, The estimated equilibrium interfacial tension values were used to plot the hexane–water (squares) and perfluorohexane–water (circles) interfacial tensions as a function of the fraction of 0.1% SDS, f_{SDS} , where the other fraction is 0.1% Zonyl. See discussion in Methods for more details.

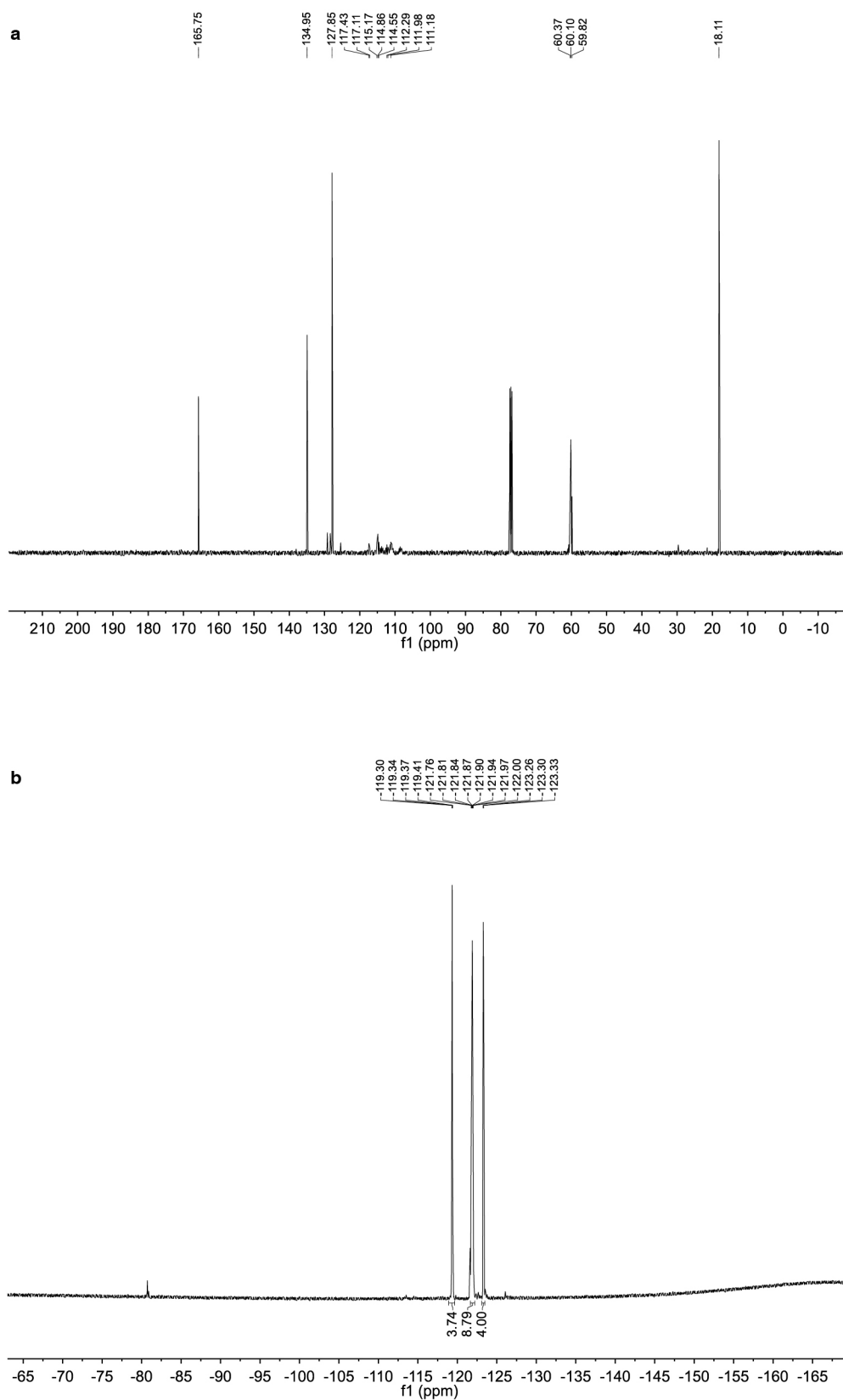


Extended Data Figure 2 | The geometry of a Janus droplet can be used to estimate the interfacial tension between hydrocarbon and fluorocarbon internal phases. **a**, Sketch of a Janus droplet consisting of hydrocarbon (grey) and fluorocarbon (white) phases within an aqueous outer phase. The radii of curvature of the H–W (R_H), F–W (R_F) and F–H (R_{FH}) interfaces are related to their respective interfacial tensions through the Young–Laplace equation. The diameter of the circle of contact between the two phases (dashed line) is denoted as D . **b**, The Janus droplet is composed of three spherical caps, and the volume, V_{cap} , of each constituent spherical cap is a function of the radius of curvature of the spherical surface and the base diameter D . Here we show the cap at the intersection of the hydrocarbon and fluorocarbon phases in

which V_{cap} is a function of R_{FH} and D . **c**, An exemplary image of a hexane–perfluorohexane Janus droplet obtained at $f_{\text{SDS}} = 0.6$, which was used to estimate R_{FH} and, in turn, γ_{FH} . **d**, The droplet pictured in **c** is subjected to edge detection to determine the H–W and F–W interfaces. **e**, The resulting edges are fitted to circles (red lines with green centres). The diameter of the circle of contact is then computed (green line). Given the ratio of the volumes of the two phases, we then determined the radius of curvature, R_{FH} , of the hexane–perfluorohexane interface, which was subsequently used to estimate γ_{FH} . See discussion in Methods for more details and Extended Data Table 2 for estimated values of γ_{FH} .



Extended Data Figure 3 | Reaction scheme and ^1H -NMR of the fluorinated crosslinker. **a**, Reaction scheme for the synthesis of the fluorinated crosslinker. BHT, 3,5-di-*tert*-butyl-4-hydroxytoluene. **b**, ^1H -NMR spectrum of the fluorinated crosslinker.



Extended Data Figure 4 | ^{13}C -NMR and ^{19}F -NMR of the fluorinated crosslinker. **a**, ^{13}C -NMR spectrum of the fluorinated crosslinker. **b**, ^{19}F -NMR of the fluorinated crosslinker.

Extended Data Table 1 | Parameter values obtained by fitting an empirical model to interfacial tension measurements obtained using pendant-drop goniometry

0.1% SDS fraction (f_{SDS})	Hexane-water interface			Perfluorohexane-water interface		
	γ_{eqb} (mN m ⁻¹)	t^* (s)	R^2	γ_{eqb} (mN m ⁻¹)	t^* (s)	R^2
0.00	7.33 ± 0.12	156 ± 24	0.925	4.42 ± 0.14	159 ± 42	0.980
0.20	7.15 ± 0.03	337 ± 15	0.989	5.80 ± 0.04	106 ± 5	0.979
0.40	7.21 ± 0.14	315 ± 53	0.956	6.49 ± 0.05	233 ± 22	0.978
0.60	6.33 ± 0.22	293 ± 113	0.947	6.45 ± 0.12	301 ± 54	0.982
0.90	6.39 ± 0.02	35 ± 1	0.992	7.66 ± 0.03	57 ± 1	0.994
0.95	5.78 ± 0.06	41 ± 1	0.993	7.73 ± 0.06	85 ± 2	0.991
1.00	14.04 ± 0.15	21 ± 3	0.909	19.56 ± 0.3	702 ± 50	0.967

Dynamic interfacial tension data were fitted to a model (Methods) to estimate equilibrium interfacial tension values. The parameters obtained from the fitting are presented here. The larger the value of the characteristic time, t^* , the slower the decay of the interfacial tension to its equilibrium value, γ_{eqb} . The value of R^2 is a measure of how well the model fits the experimental data. For all the data used in this study, R^2 was above 0.9, signifying a good fit.

Extended Data Table 2 | Values of γ_{FH} were estimated from geometrical analysis of images of Janus droplets and are independent of the composition of the surfactant added to the system, f_{SDS}

	$f_{\text{SDS}} = 0.4$	$f_{\text{SDS}} = 0.6$
γ_{FH} (mN m ⁻¹)	1.06	0.93
	1.17	1.18
	0.99	1.10

The values of γ_{FH} presented here were obtained by computing the radii of curvature of the hexane–perfluorohexane, hexane–water and perfluorohexane–water interfaces, and subsequently using the Young–Laplace equation (Methods). Six droplets in total, three each under the conditions $f_{\text{SDS}} = 0.4$ and $f_{\text{SDS}} = 0.6$, were used to estimate a value of $\gamma_{\text{FH}} = 1.07 \pm 0.1 \text{ mN m}^{-1}$.

An extremely high-altitude plume seen at Mars' morning terminator

A. Sánchez-Lavega^{1,2}, A. García Muñoz³, E. García-Melendo^{1,4}, S. Pérez-Hoyos^{1,2}, J. M. Gómez-Forrellad⁴, C. Pellier⁵, M. Delcroix⁵, M. A. López-Valverde^{2,6}, F. González-Galindo^{2,6}, W. Jaeschke⁷, D. Parker⁸, J. Phillips⁹ & D. Peach¹⁰

The Martian limb (that is, the observed 'edge' of the planet) represents a unique window into the complex atmospheric phenomena occurring there. Clouds of ice crystals (CO₂ ice or H₂O ice) have been observed numerous times by spacecraft and ground-based telescopes, showing that clouds are typically layered and always confined below an altitude of 100 kilometres; suspended dust has also been detected at altitudes up to 60 kilometres during major dust storms^{1–6}. Highly concentrated and localized patches of auroral emission controlled by magnetic field anomalies in the crust have been observed at an altitude of 130 kilometres⁷. Here we report the occurrence in March and April 2012 of two bright, extremely high-altitude plumes at the Martian terminator (the day–night boundary) at 200 to 250 kilometres or more above the surface, and thus well into the ionosphere and the exosphere^{8,9}. They were spotted at a longitude of about 195° west, a latitude of about –45° (at Terra Cimmeria), extended about 500 to 1,000 kilometres in both the north–south and east–west directions, and lasted for about 10 days. The features exhibited day-to-day variability, and were seen at the morning terminator but not at the evening limb, which indicates rapid evolution in less than 10 hours and a cyclic behaviour. We used photometric measurements to explore two possible scenarios and investigate their nature. For particles reflecting solar radiation, clouds of CO₂-ice or H₂O-ice particles with an effective radius of 0.1 micrometres are favoured over dust. Alternatively, the plume could arise from auroral emission, of a brightness more than 1,000 times that of the Earth's aurora, over a region with a strong magnetic anomaly where aurorae have previously been detected⁷. Importantly, both explanations defy our current understanding of Mars' upper atmosphere.

On 12 March 2012, amateur astronomers reported a small protrusion above the morning terminator of the Martian southern hemisphere (Fig. 1). The protrusion became more prominent over the following days, and on 20 and 21 March it was captured by at least 18 observers employing 20–40-cm telescopes at wavelengths from blue to red (~450–650 nm; Extended Data Figs 1–3). The feature was not detectable when it was passing Mars' central meridian or when it reached the opposite afternoon limb. Available images and animations produced with the MARCI (Mars Color Imager) instrument on board the Mars Reconnaissance Orbiter (see Methods) do not show the feature, because such animations are essentially mosaics formed by planet strips taken at ± 2 h of 15:00 local Martian solar time (when the limb plume is not observed)¹⁰. The 20–21 March measurements show that the feature extends from mean latitudes $-38.7^\circ \pm 6.4^\circ$ to $-49.7^\circ \pm 4.2^\circ$ (~460 km) with extremes from the mean ranging from -32.3° to -53.9° (~900 km), and when rotating into view, its longitude (westward reference system) spans $190.2^\circ \pm 4.2^\circ$ to $201.4^\circ \pm 6.8^\circ$ (~660 km) with extremes from the mean ranging from 186° to 208.2° (~1,310 km). Thus, the feature is approximately above

the southeastern area of Terra Cimmeria¹¹. The event occurred when the Martian solar heliocentric longitude was $L_s = 85^\circ$ – 90° (early winter of the southern hemisphere), and the daily mean insolation gradient at Terra Cimmeria was large¹².

The plume was detected for 11 consecutive days from 12 to 23 March. There are however no observations of the terminator above Terra Cimmeria from 24 March to 1 April (when it was not observed), a fact that prevents us from putting an end date to the plume. Remarkably, the aspect of the features changed rapidly, their shapes going from double blob protrusions to pillars or finger-plume-like morphologies (Fig. 1). On 6 April a second, similar, event was observed near the same area and lasted until at least 16 April (Extended Data Fig. 2). Each event had an overall lifetime ≥ 10 days, but changes over timescales < 12 h between dawn and dusk were accompanied by day-to-day variability. A survey of full disk Martian images taken with the Hubble Space Telescope (HST) between 1995 and 1999, and of amateur image databases (the latter come from the Association of Lunar and Planetary Observers-Japan, and Société Astronomique de France (Commission Surfaces Planétaires); Methods) from Mars apparitions between 2001 to 2014 (a total of about 3,500 images), shows the occasional presence of clouds at the limb similar to those observed by spacecraft. These clouds, however, are typically less prominent and conspicuous than in the 2012 events. Exceptionally, a set of HST images from 17 May 1997 show an abnormal plume, which is included here for photometric study^{13,14}.

We have measured the altitude of the plume top as projected onto the dark background. Only the highest spatial resolution images and those that allowed us to monitor the plume as it rotates into view from behind the terminator were considered for measurements. These were two series (24 images in total) from 20 and 21 March lasting 50 min and 70 min, respectively. In Fig. 2 we show the projected altitude of the plume top at terminator as a function of the terminator west longitude at latitude -45° (increasing longitudes with time). Measured single maximum altitudes on these two days were 200 ± 25 km for 20 March and 280 ± 25 km for 21 March. The feature's projected altitude was measured over time as the planet rotated. Fitting these altitudes to a second-order polynomial resulted in maximum plume top altitudes of 185 ± 30 km (20 March, west longitude at terminator 211.5°) and 260 ± 30 km (21 March, west longitude at terminator 214.5°). We have compared this fit to a model of the projected altitude that would result from a feature approaching the terminator and rotating into view for the Martian viewing geometric conditions (see Methods and Extended Data Fig. 4). The model fitting of the data results in a plume top altitude of 200 ± 50 km for 20 March and 260 ± 50 km for 21 March. We propose that the deviations between model fit and data for 21 March could originate from horizontal structure in the plume or from time variability over < 1 h. This analysis shows the extreme altitude of the plume top to be above

¹Departamento de Física Aplicada I, ETS Ingeniería, Universidad del País Vasco, Alameda Urquijo s/n, 48013 Bilbao, Spain. ²Unidad Asociada Grupo Ciencias Planetarias UPV/EHU-IAA (CSIC), Alameda Urquijo s/n, 48013 Bilbao, Spain. ³European Space Agency, ESTEC, 2201 AZ Noordwijk, The Netherlands. ⁴Fundació Observatori Esteve Duran, Montseny 46, 08553 Seva (Barcelona), Spain. ⁵Commission des Surfaces planétaires, Société Astronomique de France, 3, rue Beethoven, 75016 Paris, France. ⁶Instituto de Astrofísica de Andalucía, CSIC, Glorieta de la Astronomía, 3, 18008 Granada, Spain. ⁷Association of Lunar and Planetary Observers, 200 Lawrence Drive, West Chester, Pennsylvania 19380, USA. ⁸Association of Lunar and Planetary Observers, 12911 Lerida Street, Coral Gables, Florida 33156, USA. ⁹Association of Lunar and Planetary Observers, Charleston, 570 Long Point Road STE 230, Mount Pleasant, South Carolina 29464, USA. ¹⁰British Astronomical Association, Burlington House, Piccadilly, London W1J 0DU, UK.

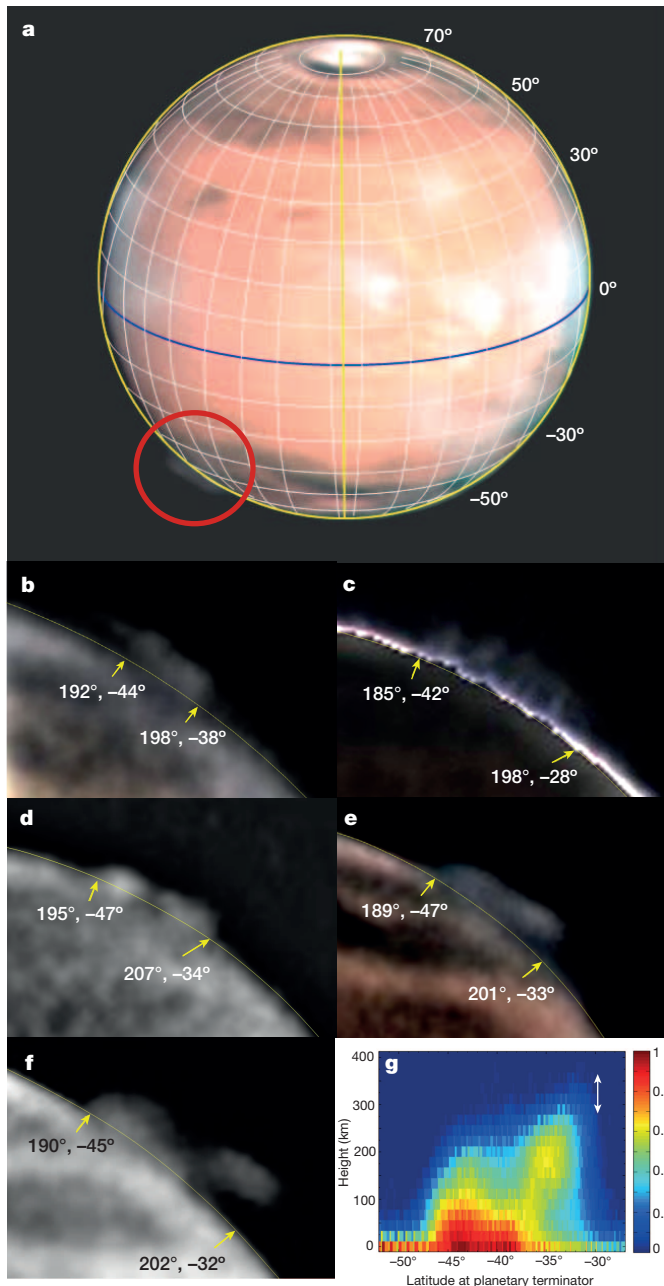


Figure 1 | A high-altitude plume at the Martian terminator. **a**, Navigated image (that is, observed planetary disk and fit to model: yellow line, disk; blue line, equator) that shows the feature location (red circle) on March 21 03:21 (image, D. Parker). Dates and times are UTC; image orientation (north, up; east, right). **b–f**, Plume images with west longitudes and latitudes at terminator (inverted orientation to facilitate visibility) for dates as follows (yellow line marks the terminator): **b**, March 20 02:45 (image W.J.); **c**, March 21 02:51 (W.J.); **d**, March 21 03:45 (J.P.); **e**, March 21 03:21, colour composite (D. Parker); **f**, March 21 03:21, red filter (D. Parker). **g**, Plume relative brightness map for March 21 03:21, red image (D. Parker). The white doubled-ended vertical arrow in **g** indicates the estimated error for altitude determination (± 50 km or 4 pixel size), which is on average 2 s.d. of the individual point uncertainty.

200 km, never before observed on Mars, and reaching the ionosphere and exosphere.

We have used the set of unprocessed images from 20 and 21 March to measure the plume's average reflectivity (I/F) at the three available wavelengths over a box of size ~ 200 km \times 200 km in the vertical and along-terminator directions (I , reflected radiation intensity; πF , incident solar flux). In order to gain insight into the photometric behaviour of

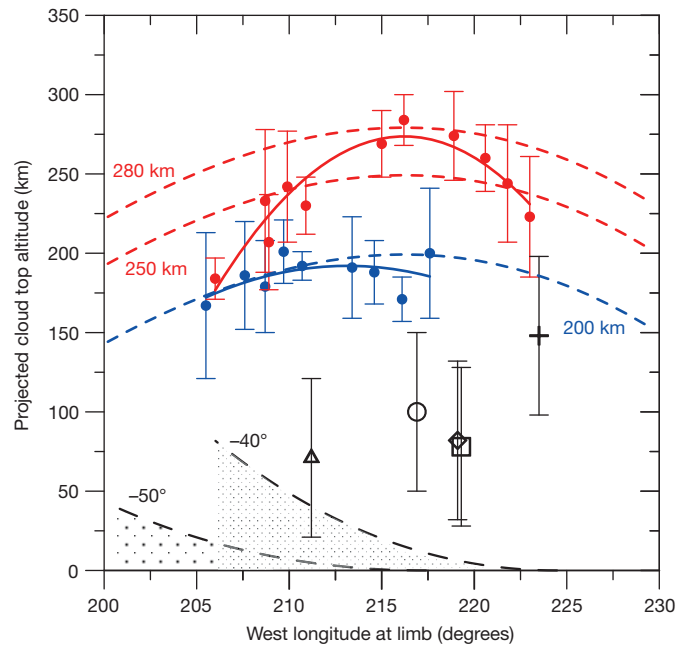


Figure 2 | Plume top altitude and its rapid changes. Filled circles show the plume's altitude when rotating into view, as given by the west longitude at the terminator (longitudes increasing with time) for March 20 02:02–02:51 UT, blue, and March 21 02:40–03:50 UT, red. Error bars, s.d. for $n = 5$ –10 measurements. Solid curves, second degree polynomial fits to these two data sets. Dashed lines, altitudes predicted by a geometric model for a feature placed at the terminator with top altitude 250 or 280 km (red) and 200 km (blue) at west longitude 216°. The single data points are individual measurements on March dates as follows (error bars, 2 s.d., $n = 5$): 12 (cross), 15 (rhomb), 17 (square), 19 (circle), 22 (triangle). Dashed lines show the projected altitude of the Mars shadow (dotted area under these curves) at two latitudes.

the plume and have a comparison reference, we determined the reflectivity from 255 to 1,042 nm wavelength of a plume captured in 1997 by the HST Wide Field Planetary Camera (Extended Data Fig. 5). The 1997 plume occurred (with its base on the terminator) at equatorial latitude $-2.9^\circ \pm 0.7^\circ$, west longitude $99.1^\circ \pm 4^\circ$ and solar longitude $L_s = 119.5^\circ$. Its plume-like shape, horizontal extent of 10° (~ 590 km) and morning occurrence suggest that both the 1997 and 2012 features are related phenomena. The actual altitude of the 1997 HST feature cannot be precisely determined because of the lack of a rotation image sequence, but we can constrain its altitude to be between 50 km (for the case of a large horizontal feature illuminated by a grazing Sun, that is, close to 0° solar illumination) and 480 km (for the case of a thin vertical feature under 90° solar illumination). Figure 3 shows the measured reflectivity (I/F) for both the 1997 and 2012 events. The spectra appear flat within the quoted uncertainties.

To interpret this phenomenon we explore two scenarios. First, we assume that the plume is a cloud formed by particles of H_2O -ice, CO_2 -ice or dust. We performed forward radiative transfer modelling of the reflected intensity and compared it with the observations. The model considers multiple scattering and curvature effects for grazing illumination and a slanted view (see Methods). For the optical properties, we assumed Mie theory¹⁵ and wavelength-dependent indices of refraction^{16,17}. As shown in Fig. 3, dust particles can be ruled out for the 1997 event and show only marginal agreement with the 2012 data (see also Extended Data Fig. 6). Typically, the best fits occur for CO_2 or H_2O ice particles with effective radii of $0.1^{+0.1}_{-0.004}$ μm , which is consistent with the particle size of mesospheric clouds observed at night⁴. Comparably good fits are obtained for a broad range of particle effective variances (0.1–2.0), which means that the latter parameter cannot be constrained with the available information. The best fit occurs for a layer of vertical thickness 100 km and a nadir optical depth $\tau_N > 0.5$ for both the 1997 and

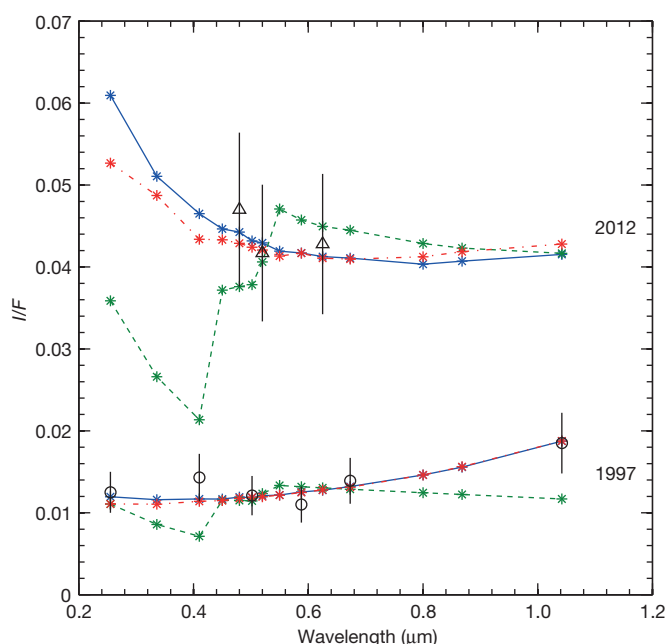


Figure 3 | Plume reflectivity and radiative transfer model comparison. Spectral reflectivity of the 20–21 March 2012 event (triangles) and of the 17 May 1997 event (circles) from ground-based and HST observations, respectively. The error bars represent the average quadratic deviation of the measured reflectivity in the integration box. Observations are compared with the best-fitting model for the case where the protrusion is assumed to be a cloud formed by spherical particles: CO₂-ice (blue solid line), H₂O-ice (red dot-dashed line) or Martian dust (green dashed line). Stars indicate the wavelength grid used for radiative transfer computations.

2012 events, which implies number densities of 0.01 particles per cm³ (see Methods and Extended Data Figs 7 and 8). Vaporization of these particles will produce an increase in the gas concentration by a factor of $\sim 1,000$ for H₂O and 5% for CO₂ (see Methods).

According to a general circulation model (GCM)^{18,19} for conditions specific to the observation dates, H₂O condensation at the relevant altitudes requires either anomalously cold thermospheric temperatures (with temperature drop >50 K) or an unusual increase in the H₂O mixing ratio from 10^{-4} to complete saturation above 140 km (Fig. 4)²⁰. CO₂ condensation would require an even larger temperature drop of 100 K above 125 km. On the other hand, explaining the cloud as formed by dust would require vigorous vertical transport up to at least 180 km above the surface (but this is not predicted by GCMs) or by vigorous updrafts due to dry convection under high insolation (dust heating follows from the absorption of solar radiation)^{21,22}. Thus, upward motions are more likely to occur at noon than in the morning; all of which, together with the photometric data, makes the dust hypothesis difficult to support.

In the second scenario, we explored whether the 2012 plume might be attributable to an aurora. Mars aurorae have been observed near where the plume occurs, a region with a large anomaly in the crustal magnetic field (at 175° west)⁷ that can drive the precipitation of solar wind particles into the atmosphere^{7,23}. The meridional extent of the plume (~ 500 km, both consistent with aurorae⁷) and its variability could also support this hypothesis. Mars' ultraviolet aurora is dominated by the Cameron bands of carbon monoxide (CO), with limb intensities of kilorayleighs (kR; refs 7, 24); visible counterpart emissions are also expected. However, quantitative estimates of the aurora intensity defy such a hypothesis. An I/F value of 0.04 at 550 nm (Fig. 3) translates into an auroral limb intensity of 3,600 megarayleighs (MR), much more than the 1-MR nadir emission of strong terrestrial aurorae²⁵. Furthermore, an auroral hypothesis requires an exceptional influx of energetic particles over days, although solar activity in March 2012 was not unusually high²⁶. Even accounting for deficiencies in the understanding of auroral

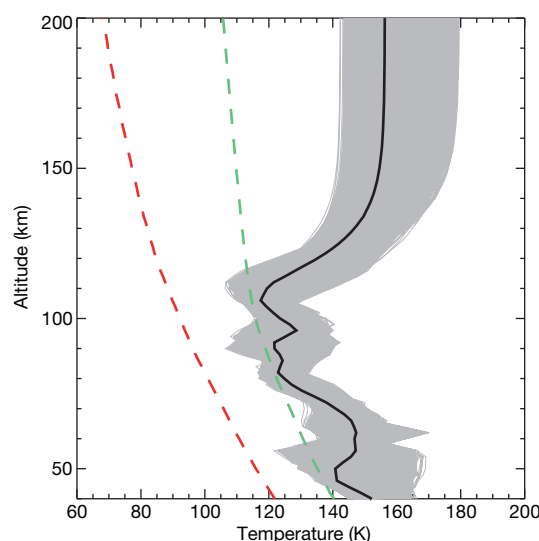


Figure 4 | Atmospheric temperature profile, and water and carbon dioxide condensation temperatures. These are calculated for the conditions of the 2012 event (latitude -40° , longitude 200° , $L_s = 90^\circ$, local time LT = 6 h) according to a general circulation model^{18,19}. Atmospheric temperature is shown by the black line, with grey profiles indicating values extending the range by $\pm 10^\circ$ for latitude, 10° in longitude and 2 h in LT. Condensation temperatures for water (dashed green line) and carbon dioxide (dashed red line) are obtained from the saturation vapour pressure curves of both compounds²⁰.

excitation^{27,28}, the extrapolation from ultraviolet intensities to the visible falls short of the reported plume brightness by orders of magnitude. Confirmation or rejection of the auroral hypothesis is, however, feasible merely by extended monitoring from ground or space.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 August; accepted 17 December 2014.

Published online 16 February 2015.

- Määttä, A. *et al.* A complete climatology of the aerosol vertical distribution on Mars from MEX/SPICAM UV solar occultations. *Icarus* **223**, 892–941 (2013).
- Heavens, N. G. *et al.* The vertical distribution of dust in the Martian atmosphere during northern spring and summer: observations by the Mars Climate Sounder and analysis of zonal average vertical dust profiles. *J. Geophys. Res. Planets* **116**, E04003 (2011).
- Määttä, A. *et al.* Mapping the mesospheric CO₂ clouds on Mars: MEX/OMEGA and MEX/HRSC observations and challenges for atmospheric models. *Icarus* **209**, 452–469 (2010).
- Montmessin, F. *et al.* Hyperspectral imaging of convective CO₂ ice clouds in the equatorial mesosphere of Mars. *J. Geophys. Res.* **112**, E11S90 (2007).
- Montmessin, F. *et al.* Subvisible CO₂ ice clouds detected in the mesosphere of Mars. *Icarus* **183**, 403–410 (2006).
- Clancy, R. T. *et al.* CRISM limb observations of Mars mesospheric ice clouds: two new results. *Proc. Eighth Int. Conf. on Mars*, abstr. 1006, <http://www.hou.usra.edu/meetings/8thmars2014/pdf/program.pdf> (2014).
- Bertaux, J. L. *et al.* Discovery of an aurora on Mars. *Nature* **435**, 790–794 (2005).
- Strobel, D. F. in *Atmospheres in the Solar System: Comparative Aeronomy* (eds Mendillo, M., Nagy, A. & Waite, J. H.) 39–54, 191–202 (American Geophysical Union, 2002).
- Yagi, M. *et al.* Mars exospheric thermal and non-thermal components: seasonal and local variations. *Icarus* **221**, 682–693 (2012).
- Bell, J. F., III *et al.* Mars Reconnaissance Orbiter Mars Color Imager (MARCI): instrument description, calibration, and performance. *J. Geophys. Res.* **114**, E08S92 (2009).
- Zuber, M. T. *et al.* The Mars Observer Laser Altimeter investigation. *J. Geophys. Res.* **97**, 7781–7797 (1992).
- Córdoba-Jabonero, C., Zorzano, M. P., Selsis, F., Patel, M. R. & Cockell, C. S. Radiative habitable zones in Martian polar environments. *Icarus* **175**, 360–371 (2005).
- Minami, M. *Communication in Mars Observations*. No. 283 <http://www.kwasan.kyoto-u.ac.jp/~cmo/cmomn2/2830AA/index.htm> (2003).
- Pellier, C. Martian terminator projections observed by the HST. In *Communications in Mars Observations* No. 400, <http://www.kwasan.kyoto-u.ac.jp/~cmo/cmomn4/CMO400.pdf> (2012).

15. Mishchenko, M. I., Travis, L. D. & Lacis, A. A. *Scattering, Absorption and Emission of Light by Small Particles* (Cambridge Univ. Press, 2002).
16. Warren, S. G. Optical constants of carbon dioxide ice. *Appl. Opt.* **25**, 2650–2674 (1986).
17. Fedorova, A. A. *et al.* Evidence for a bimodal size distribution for the suspended aerosol particles on Mars. *Icarus* **231**, 239–260 (2014).
18. González-Galindo, F., Määttänen, A., Forget, F. & Spiga, A. The Martian mesosphere as revealed by CO₂ cloud observations and general circulation modeling. *Icarus* **216**, 10–22 (2011).
19. González-Galindo, F. *et al.* Three-dimensional Martian ionosphere model: I. The photochemical ionosphere below 180 km. *J. Geophys. Res.* **118**, 2105–2123 (2013).
20. Sánchez-Lavega, A., Pérez-Hoyos, S. & Hueso, R. Condensate clouds in planetary atmospheres: a useful application of the Clausius-Clapeyron equation. *Am. J. Phys.* **72**, 767–774 (2004).
21. Heavens, N. G. *et al.* Vertical distribution of dust in the Martian atmosphere during northern spring and summer: High-altitude tropical dust maximum at northern summer solstice. *J. Geophys. Res. Planets* **116**, E01007 (2011).
22. Spiga, A., Faure, J., Madeleine, J. P., Määttänen, A. & Forget, F. Rocket dust storms and detached dust layers in the Martian atmosphere. *J. Geophys. Res. Planets* **118**, 746–767 (2013).
23. Mitchell, D. L. *et al.* A global map of Mars' crustal magnetic field based on electron reflectometry. *J. Geophys. Res.* **112**, E01002 (2007).
24. Leblanc, F. *et al.* Observations of aurorae by SPICAM ultraviolet spectrograph on board Mars Express: simultaneous ASPERA-3 and MARSIS measurements. *J. Geophys. Res.* **113**, A08311 (2008).
25. Chamberlain, J. W. *Physics of the Aurora and Airglow*, Ch. 4, 5 (American Geophysical Union, 2013).
26. SpaceWeatherLife.com <http://www.spaceweatherlive.com/en/archive/2012/03/12/rsga> (2014).
27. Fox, J. The ionospheric source of the red and green lines of atomic oxygen in the Venus nightglow. *Icarus* **221**, 787–799 (2012).
28. Slinger, T. G. *et al.* Photoemission phenomena in the Solar System. *Space Sci. Rev.* **139**, 267–310 (2008).

Acknowledgements This work was supported by the Spanish MINECO project AYA2012-36666, FEDER, Grupos Gobierno Vasco IT765-13 and UPV/EHU UFI11/55. The IAA (CSIC) team was supported by the Spanish MINECO through the CONSOLIDER programme ASTROMOL CSD2009-00038 and AYA2011-30613-CO2-1. F.G.-G. is funded by a CSIC JAE-Doc contract co-financed by the European Social Fund.

Author Contributions A.S.-L. coordinated the study and performed plume measurements, photometric calibration and participated in the models study; A.G.M. studied aurora and together with S.P.-H. performed radiative transfer modelling; E.G.-M. performed with A.S.L. the geometric modelling; M.A.L.-V. and F.G.-G. performed the GCM model calculations and aurora studies, and F.G.-G. performed the evaporation calculations; J.M.G.-F., C.P. and M.D. performed plume measurements; W.J., D. Parker, J.P. and D. Peach performed ground-based images. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.S.-L. (agustin.sanchez@ehu.es).

METHODS

Image availability. *Amateur images.* Available at the following databases: Association of Lunar and Planetary Observers-Japan, <http://alpo-j.asahikawa-med.ac.jp/Latest/Mars.htm> (2012), and Société Astronomique de France (Commission surfaces Planétaires), <http://www.astrosurf.com/pellier/mars> (2012).

Hubble Space Telescope images. Available at NASA Planetary Data System (PDS) or at ESA HST archive, <http://archives.esac.esa.int/hst/#Q-R-S+D-1> (2012).

MRO MARCI Weather Reports. Malin Space Science Systems. Captioned Image Release No. MSSS-216 – 21 March 2012 http://www.msss.com/msss_images/2012/03/21/ and Captioned Image Release No. MSSS-217 – 28 March 2012 http://www.msss.com/msss_images/2012/03/28/.

MOLA maps. Available at http://mola.gsfc.nasa.gov/images/topo_labeled.jpg.

Image analysis and measurement. Most images were obtained with broadband filters covering the blue (B, $\lambda_{\text{eff}} = 450 \pm 50$ nm), green (G, $\lambda_{\text{eff}} = 520 \pm 50$ nm) and red (R, $\lambda_{\text{eff}} = 625 \pm 75$ nm) spectral bands. Each image was constructed by summing, aligning and re-centring sequences of frames captured in video mode, and application of the 'lucky imaging' method^{29–31} and the 'wavelet' as the main processing technique³⁰. The Airy disk for a telescope of 356 mm in diameter (Fig. 1a, b, d–g and photometric analysis) at a wavelength of 450 nm is 0.317 arcsec and the plate scale employed was 0.045 arcsec per pixel equivalent to an effective pixel size of 25 km on Mars, typically used in high resolution planetary images where recording features depends, for excellent seeing conditions, on resolution and features contrast³². The total spatial coverage of the projected plume is ~ 850 pixels (21 March, Fig. 1g). Image navigation, that is, the determination of the planetary limb and terminator was performed using two well-tested software packages (LAIA, WinJupos; ref. 33 and <http://www.grischa-hahn.homepage.t-online.de/astro/index.htm>). The bright Martian limb was used as a reference but not the terminator whose definition is affected by the presence of the dark and bright features. Navigation was accurately controlled by measuring the position of well-known surface features (Olympus Mons and Tharsis volcanoes). The projected plume top altitude on the sky plane (defined as the border between the feature and the dark background, contrast forced) was simply determined from its distance to the Mars disk centre and subtraction of the measured Martian radius (all in pixel units). Navigation and image measurements were performed independently by four of us. From these multiple measurements, we derived the mean values of the Mars radius and top altitude for each image. The uncertainties represent the root mean square (r.m.s.) values of the altitude determinations. We estimate that the positioning error is ± 2 pixels and that, for the above plate-scale, this corresponds to ± 50 km. Uncertainties in plume measurements result from brightness diffusion at the top and sides of the plumes and from the determination of the base location of the plume relative to the navigated terminator, including its brightness irregularities.

As a further step in navigation control, we used the image series of 20 March by W.J. to compare the results of the top altitude measurements by two different procedures: (1) through measurements on each image of the planet and plume top radius to calculate the plume maximum altitude; (2) through determination of the mean radius for the whole image series (and its uncertainty) and then using this mean value for each image to determine the top altitude. Both approaches yield consistent results within the uncertainties indicated in the text.

The measured top altitude values (H) and the terminator west longitudes (L , defined as the central meridian longitude for each observing time $+70^\circ$) for 20 and 21 March, fitted to a second order polynomial $H = a_0 + a_1 L + a_2 L^2$, gave the peak values at $L_{\text{max}} = -a_1/2a_2$, and then the corresponding H_{max} . The retrieved coefficients are, for 20 March, $a_0 = -15,233.8$, $a_1 = 144.7$, $a_2 = -0.3395$, and for 21 March, $a_0 = -43,164.8$, $a_1 = 401.8$, $a_2 = -0.929$.

Plume altitude model. *Angular distance between the plume's base and the central meridian.* Our coordinate system is placed at the centre of Mars: the x axis will be on the observer's visual and the y - z plane is coincident with the sky plane. Coordinate system geometry is depicted in Extended Data Fig. 4, with the cloud rising near the west morning limb and the sub-Earth point on the equator. The longitude of the Central Meridian (CM) is λ_{CM} , and the latitude and longitude of the plume are respectively φ_c and λ_c . The longitudinal angular distance of the plume to the CM is related to the angular distance β between the plume and the limb taking as a reference the negative y axis, by:

$$90^\circ - \beta = \lambda_c - \lambda_{\text{CM}} \quad (1)$$

where β is positive if the limb is behind the cloud, and negative if the cloud is located behind the limb.

Limit altitude for plume visibility. Let $z(\beta)$ be the visible projected plume altitude above Mars' limb not blocked by the planet's disk when $\beta < 0$. When the plume is behind the limb it protrudes from the planet's limb only if its total projected length on the y - z plane from the planet centre is larger than the planet's radius R_M . With

c_y and c_z being respectively the y and z coordinates of the cloud top, we have $\{c_y^2 + c_z^2\}^{1/2} = R_M + z(\beta)$. In terms of β and the real cloud altitude H , we get

$$\begin{cases} c_y = -(R_M + H) \cos \varphi_c \cos \beta \\ c_z = (R_M + H) \sin \varphi_c \end{cases} \quad (2)$$

The minimum height of the plume H_{min} to protrude from the limb ($z(\beta) = 0$) is

$$H_{\text{min}} = \frac{R_M}{\{1 - \cos^2 \varphi_c \sin^2 \beta\}^{1/2}} - R_M \quad (3)$$

Real plume altitude derived from its projected length. When $\beta < 0$, by expression (2), $z(\beta)$ must satisfy

$$z(\beta) = (R_M + H) \{ \cos^2 \varphi_c \cos^2 \beta + \sin^2 \varphi_c \}^{1/2} - R_M \quad (4)$$

and the actual plume altitude is:

$$H = \frac{z(\beta) + R_M}{\{1 - \cos^2 \varphi_c \sin^2 \beta\}^{1/2}} - R_M \quad (5)$$

If $\beta > 0$, the whole plume is visible (we see its projected total height). The cloud height H is now given by:

$$H = \frac{z(\beta)}{\{1 - \cos^2 \varphi_c \sin^2 \beta\}^{1/2}} \quad (6)$$

Correction when Mars has a declination D_E as observed from Earth. By taking the declination angle $+D_E$ when the north pole is tilted towards Earth and $-D_E$ when the planet's south pole is visible, the transformation matrix due to a rotation around the y axis by an angle $+D_E$ is

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos D_E & 0 & \sin D_E \\ 0 & 1 & 0 \\ -\sin D_E & 0 & \cos D_E \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (7)$$

Therefore, expressions (5) and (6) must be corrected taking into account that:

$$\begin{cases} y' = y \\ z' = -x \sin D_E + z \cos D_E \end{cases} \quad (8)$$

In the rotated coordinate system, the projected plume altitude is $z'(\beta) = \{c_y'^2 + c_z'^2\}^{1/2} - R_M$, with c_y' and c_z' being the transformed coordinates of the plume top, with the non-transformed coordinates of the plume top given by:

$$\begin{aligned} c_x &= (R_M + H) \cos \varphi_c \sin \beta \\ c_y &= -(R_M + H) \cos \varphi_c \cos \beta \\ c_z &= (R_M + H) \sin \varphi_c \end{aligned} \quad (9)$$

Expressing c_y' and c_z' in terms of c_x , c_y and c_z when the plume is behind the limb gives its real height

$$H = \frac{z'(\beta) + R_M}{\{\cos^2 \varphi_c \cos^2 \beta + (\sin \varphi_c \cos D_E - \cos \varphi_c \sin \beta \sin D_E)^2\}^{1/2}} - R_M \quad (10)$$

Owing to the planet's tilt, the plume will be on the limb not for $\beta = 0$ but for a different value β_{lim} . Therefore, for $\beta = \beta_{\text{lim}}$, $z'(\beta) = 0$ if $H = 0$. Applying this condition to expression (10) and after some algebraic manipulations, we obtain:

$$\beta_{\text{lim}} = -\sin^{-1} \tan(\varphi_c) \times \tan(D_E) \quad (11)$$

When $D_E = 0$ we recover $\beta_{\text{lim}} = 0$. Note that at the equator β_{lim} is always 0 regardless the value of D_E . Expression (10) is valid for $\beta \leq \beta_{\text{lim}}$ (the plume is behind the planet's limb). When $\beta \geq \beta_{\text{lim}}$ the plume is before the limb and is not blocked by the planet and we have:

$$H = \frac{z'(\beta)}{\{\cos^2 \varphi_c \cos^2 \beta + (\sin \varphi_c \cos D_E - \cos \varphi_c \sin \beta \sin D_E)^2\}^{1/2}} \quad (12)$$

Mars' shadow projection on the plume. Given the coordinates of the subsolar point on Mars (λ_s , φ_s) and the plume coordinates (λ_c , φ_c) we have both points separated by an angle:

$$\gamma = \cos^{-1} [\cos \varphi_s \cos \varphi_c \cos(\lambda_s - \lambda_c) + \sin \varphi_s \sin \varphi_c] \quad (13)$$

This expression is also valid when λ_s and λ_c are the respective distances to the CM. In this case, $\lambda_c = (\pi/2) - \beta$ when we are close to the morning limb and:

$$\gamma = \cos^{-1} [\cos \varphi_s \cos \varphi_c \sin(\lambda_s + \beta) + \sin \varphi_s \sin \varphi_c] \quad (14)$$

Looking at Extended Data Fig. 5, the angle $\alpha = \gamma - (\pi/2)$ and the height of the planet's shadow projected on the plume is:

$$S_C = \frac{R_M}{\cos \alpha} - R_M \quad (15)$$

For $\beta < 0$, using equations (5) and (15), the visible projected length of the planet's shadow projected on the protruding plume is:

$$S(\beta) = \frac{R_M}{\cos \alpha} \{ \cos^2 \varphi_c \cos^2 \beta + \sin^2 \varphi_c \}^{1/2} - R_M \quad (16)$$

For $\beta > 0$ the planet's shadow projected on the plume is completely visible, since the planetary disk does not block it. The real height of the shadow will be S_C again, and according to equation (6) the projected shadow's height will be:

$$S(\beta) = S_C \{ \cos^2 \varphi_c \cos^2 \beta + \sin^2 \varphi_c \}^{1/2} \\ = R_M \left[\frac{1}{\cos \alpha} - 1 \right] \{ \cos^2 \varphi_c \cos^2 \beta + \sin^2 \varphi_c \}^{1/2} \quad (17)$$

Again, if the planet's rotation axis is tilted by $\pm D_E$ degrees, we obtain

$$S(\beta) = \frac{R_M}{\cos \alpha} \{ \cos^2 \varphi_c \cos^2 \beta + (\sin \varphi_c \cos D_E - \cos \varphi_c \sin \beta \sin D_E)^2 \}^{1/2} - R_M \quad (18)$$

for $\beta \leq \beta_{\text{lim}}$, and

$$S(\beta) = S_C \{ \cos^2 \varphi_c \cos^2 \beta + (\sin \varphi_c \cos D_E - \cos \varphi_c \sin \beta \sin D_E)^2 \}^{1/2} \quad (19)$$

for $\beta \geq \beta_{\text{lim}}$.

The model predictions are shown in Fig. 2. A χ^2 calculation between the measured top altitudes considering the individual error and the predictions from the geometric model for a uniform plume shows that for 20 March $\chi^2 < 4$ for top altitudes 190–195 km (best fit), but for 21 March $\chi^2(\text{min}) \approx 20$ for top altitudes 240–250 km. The large deviations between the model and the observations for 21 March data rule out a uniform structure for the plume.

Altitude of the 1997 HST feature. The 17 May 1997 event was captured by HST on a single image at each specific wavelength, and thus they do not show the plume in rotation at the limb/terminator. Therefore we can only constrain the maximum and minimum altitudes for the feature as seen projected on the terminator. The feature protrudes on the terminator by 10° (angle α in Extended Data Fig. 4). For solar zenith angle SZA = 90° (grazing illumination) the minimum altitude is given by equation (15): $H_{\text{min}} = S_C = 53$ km. On the other hand, the maximum altitude H_{max} will occur for a vertical feature whose projection L as seen on the terminator from Earth (phase angle $\phi = 35.5^\circ$) is given by $L = \sqrt{(R_M + S_C)^2 - R_M^2}$ and then $H_{\text{max}} = L \tan \phi = 480$ km.

Photometric calibration. (1) *Ground-based observations (20–21 March 2012).* Stacked and aligned unprocessed images were used for reflectivity measurements. Because of their higher resolution, only the series B, G, R filtered images from W.J. and D. Parker were used for that purpose. The photometry from the brightest part of the cloud was compared to different bright and dark regions of Mars whose absolute (I/F) reflectivity is known³⁴. The resulting I/F can be seen as an average for the cloud at B, G and R, and from independent measurements by the team members of the same images, we estimate an uncertainty of 20%.

(2) *HST calibration (17 May 1997).* Hubble Space Telescope observations acquired in filters from the ultraviolet (F255W) to the near-infrared (F1042M) (Extended Data Fig. 5) were photometrically calibrated following the WFPC2 handbook instructions. Radiance were converted into absolute reflectivity I/F using the solar spectrum^{35,36}. The resulting I/F as a function of planetary geographical coordinates was confirmed against values given by other authors for selected locations of the planet³⁴ and with global albedo values of the planet³⁷. The intensity calibration method for ground-based images was validated with a similar procedure to that used in HST images and comparison with the above procedure.

Radiative transfer model. We conducted radiative transfer calculations of the reflectivity I/F with a backward Monte Carlo model designed for spherical-shell atmospheres³⁸. The model produces as output the full Stokes vector of radiances, but only the first vector element (intensity) was considered for comparison with the observations. The model naturally accounts for curvature effects in the (nearly) grazing illumination/viewing geometries of both the 1997 and 2012 events. In its specific implementation here, the model assumes a uniform cloud of geometrical thickness D that enshrouds the entire planet. The observer looks upon the planet terminator with a line of sight that sequentially probes the full range of altitudes where the cloud is present. For the 1997 event, plane π_{sun} (formed by the local vertical direction at the planet's terminator and the Sun illumination direction), and plane π_{sobs} (formed by the same local vertical direction and the observer viewing direction) are offset by an azimuth angle of 15° . The cosine of the polar angle between the observer's viewing direction and the local vertical at the terminator is

0.514 ± 0.02 , and is 0.404 ± 0.03 for the farthest point of the plume from the terminator. The phase angle (Sun–Mars–Earth) was 35.6° . Correspondingly, for the 2012 event, the prescribed azimuth angle and the cosine of the polar angle are 0° and 0.014 ± 0.03 , and the phase angle (Sun–Mars–Earth) was 7.15° . Solar photons enter the atmosphere following grazing trajectories, and penetrate half way through the limb before being scattered towards the observer at the terminator. To compare with the box-averaged I/F s from observations, the model-calculated I/F s were also averaged over the cloud geometrical thickness D from the observer's vantage point. A number of photon realizations between 10^4 and 10^5 ensured accuracies typically better than 5%, which is much less than the measured uncertainties. The atmospheric inputs to the model include the cloud particles' optical properties and the cloud (vertical) optical thickness τ_N , in addition to the geometrical thickness of the cloud D .

In the prescription of particle sizes, we used a power law distribution described by the two moments r_{eff} (effective radius) and v_{eff} (effective variance)¹⁵. Particle candidates are CO_2 -ice, H_2O -ice and dust, and we considered their wavelength-dependent refractive indices^{13,14}.

The particles' optical properties required in the radiative transfer calculations are ω , $p(\theta)$ and σ , which stand for the particles' single scattering albedo, scattering phase function and extinction cross-section, respectively. Each of these properties is specified at the appropriate radiation wavelength λ . For the determination of the optical properties we used Mie theory, which applies to scattering by spherical particles. If the particles are non-spherical, Mie theory may bias the properties of inferred particles that fit best the measured I/F reflectivities. The differences in optical properties between spherical and non-spherical particles depend on the particles' shape and size distribution, as well as on their composition and on the radiation wavelength. Numerical investigations have explored these differences over a limited range of parameter space¹⁵. Without additional information on the cloud particles, it is difficult to assess the consistency in the treatment with Mie theory of the radiative transfer problem. We can nevertheless tentatively quantify it by comparison against specific solutions from the literature. For $r_{\text{eff}} = 0.1 \mu\text{m}$ and $\lambda = 0.55 \mu\text{m}$, the effective size parameter $x_{\text{eff}} = 2\pi r_{\text{eff}}/\lambda \approx 1$. Focusing on the scattering phase function, which is particularly sensitive to the particles' shape, a few studies¹⁵ suggest that for $x_{\text{eff}} \approx 1$ the differences between spherical and non-spherical particles are not usually more than a few tens of per cent. These differences are less than for larger values of x_{eff} .

Overall, each radiative transfer calculation requires user-inputted values for λ , particle composition (and therefore wavelength-dependent refractive indices), r_{eff} , v_{eff} and τ_N (also dependent on wavelength). The problem of comparing reflectivities I/F from model calculations and observations is therefore multi-parametric. In the exploration of the space of parameters, we built a grid of solutions to the radiative transfer problem for D (10, 50, 100, 200 km), r_{eff} (0.01 to 2 μm ; 20 values with varying steps between 0.01 and 0.1 μm), v_{eff} (0.1, 0.2, 0.5, 1 and 2; range consistent with values adopted in prior Mars investigations^{4,15,39}) and $\tau_{n,502\text{nm}}$ ($=\tau_n$ at 502 nm; from 16 to 10^{-3} ; 31 values in steps following a power law). The total number of radiative transfer calculations for each of the 1997 and 2012 events exceeds 10^5 .

For each possible spectrum within the grid of calculations with identical particle composition, r_{eff} , v_{eff} and $\tau_{n,502\text{nm}}$, we calculated the average quadratic deviation of the model from the observed data weighted by the observation errors (χ^2). A subset of the free parameter space with best-fitting values is shown in Extended Data Fig. 6 for the 1997 event and Extended Data Fig. 7 for the 2012 event. In the former case, dust particles do not fit the data with $\chi^2 < 1.0$. While a number of sub-optimal solutions for different combinations of the free parameters can be found, the highest density of low χ^2 solutions can be interpreted as the most likely combination of the free parameters. This happens for particles with effective radii of the order of 0.1 μm , depending on the particles' effective variance and composition. In this framework, there is no way to discriminate between the variances and this also impedes making a choice between H_2O and CO_2 as the most likely particle type. Similarly, the χ^2 results for the 2012 observations show a broad range of particle sizes and optical thicknesses that provide $\chi^2 < 1.0$. They include dust particles that, in any case, result in fits that are worse than for the H_2O and CO_2 compositions. Extended Data Fig. 8 shows how, given the narrow spectral range covered during the 2012 event, it is possible to retrieve acceptable fits for particle sizes ranging from 0.1 to 1.0 μm , provided that the rest of the free parameters are adequately tuned. Observations at shorter/longer wavelengths might potentially break this degeneracy.

All in all, the radiative transfer calculations for both the 1997 and 2012 events constrain to some extent the effective particle radius r_{eff} , and the vertical extension of the cloud D , and give a lower limit for the normal optical thickness $\tau_{n,502\text{nm}}$. Unfortunately, additional data are required to constrain other parameters, such as v_{eff} or the cloud particle composition. The optical thickness $\tau_{n,502\text{nm}}$ retrieved from the model fit translates into particle density after dividing the optical thickness by the particle cross-section at 502 nm and the cloud geometrical thickness.

Code availability. All radiative transport calculations were conducted with a novel pre-conditioned backward Monte Carlo (PBMC) model³⁸. A plane-parallel version of the model is freely available through: <http://dx.doi.org/10.1051/0004-6361/201424042>. This version contains instructions to run the model, test cases and model outputs that can be compared to calculations tabulated in ref. 38. In the analysis of the Mars plume, we used a version of the PBMC suitable for spherical shell atmospheres. The spherical-shell model is available on request from A.G.M. (tonhimgm@gmail.com).

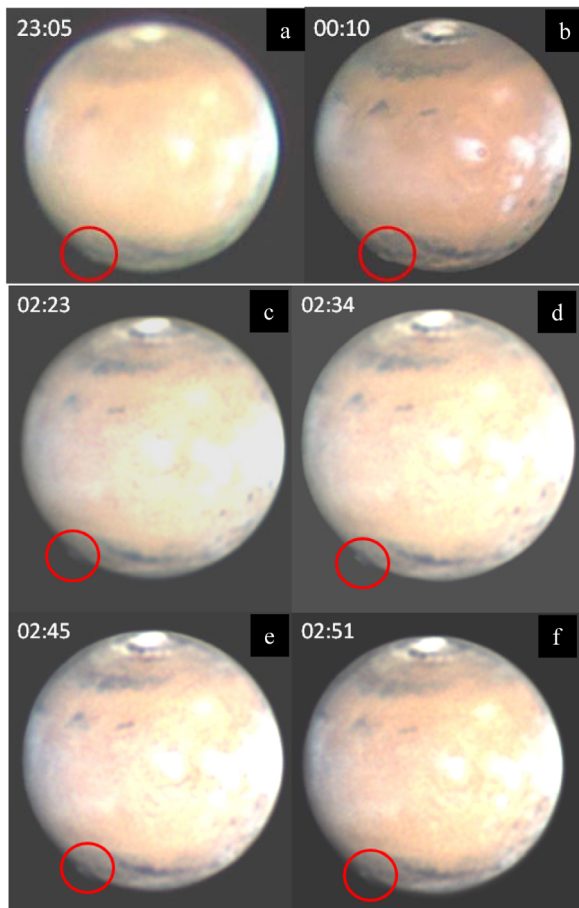
Mars General Circulation Model. As supplied in the Mars Climate Database at <http://www-mars.lmd.jussieu.fr/>.

Cloud evaporation. *H₂O-cloud.* Taking water-ice density as 0.93 g cm^{-3} (ref. 40), the mass of a spherical particle with radius $0.1 \mu\text{m}$ is $m = 4 \times 10^{-15} \text{ g}$; the corresponding number of molecules resulting from evaporation of this particle is $mN_{\text{AvO}}/\mu \approx 10^8$, taking $N_{\text{AvO}} = 6.023 \times 10^{23}$ particles per mol and the molecular weight of water is $\mu = 18 \text{ g mol}^{-1}$. For the retrieved cloud density 0.01 cm^{-3} , evaporation of the cloud will produce $\sim 10^6$ particles per cm^3 . On the other hand, the atmospheric density is $\rho = 5 \times 10^{-15} \text{ g cm}^{-3}$ at an altitude of 200 km, at latitude -40° , $L_s = 90^\circ$, 210° W longitude and local time 6 h (ref. 18). The mean molecular weight of the atmosphere at this altitude is now $\bar{\mu} = 17 \text{ g mol}^{-1}$ because of oxygen dominance with a volume mixing ratio $\text{vmr} = 0.6$ and a 13% contribution from N_2 and CO. The vmr of H_2O at this altitude is 4×10^{-6} . The number density of molecules in the atmosphere is $N(\text{H}_2\text{O}, \text{atm.}) = (\rho_{\text{H}_2\text{O}} \times N_{\text{AvO}}/\bar{\mu}) \times \text{vmr}_{\text{H}_2\text{O}} \approx 10^3 \text{ particles cm}^{-3}$. Therefore, evaporation of the water ice will produce an enrichment of water by a factor $\sim 1,000$ over normal conditions.

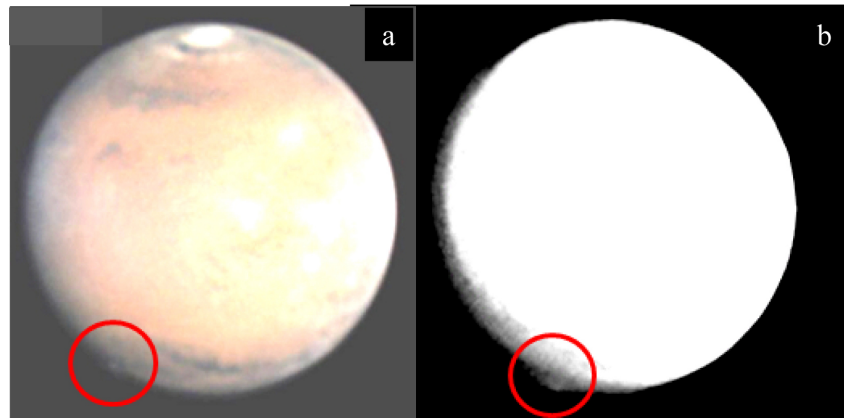
CO₂-cloud. Taking CO₂-ice density as 1.5 g cm^{-3} , the mass of a spherical particle with radius $0.1 \mu\text{m}$ will be $m = 6 \times 10^{-15} \text{ g}$ and the corresponding number of molecules resulting from evaporation is $mN_{\text{AvO}}/\mu' \approx 10^8$ where $\mu' = 44 \text{ g mol}^{-1}$ is the

molecular weight of CO₂. For the retrieved cloud particle density of 0.01 cm^{-3} , cloud evaporation will produce $\sim 10^6$ particles per cm^3 . On the other hand, from the above atmospheric conditions, and $\text{vmr CO}_2 = 0.09$, the number density of molecules in the atmosphere is $N(\text{CO}_2, \text{atm.}) = (\rho_{\text{CO}_2} \times N_{\text{AvO}}/\bar{\mu}) \times \text{vmr}_{\text{CO}_2} \approx 10^7 \text{ particles cm}^{-3}$. Therefore, CO₂ ice evaporation will produce a 5% enrichment of CO₂.

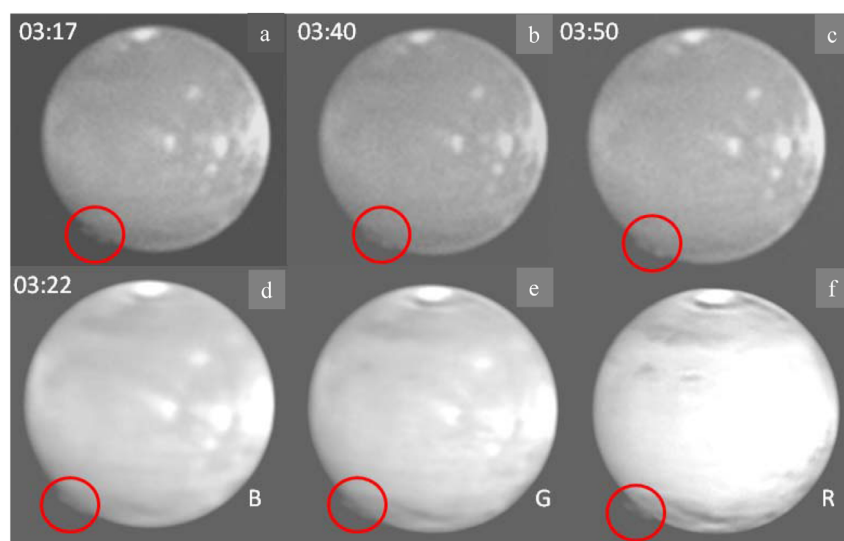
29. Law, N. M., Mackay, C. D. & Baldwin, J. E. Lucky imaging: high angular resolution imaging in the visible from the ground. *Astron. Astrophys.* **446**, 739–745 (2006).
30. Berrevoets, C. et al. RegiStax6. <http://www.astronomie.be/registax/> (2011).
31. Kraaikamp, E. AustoStakkert!2. <http://www.autostakkert.com/> (2012).
32. Peach, D. A. The 2010 perihelic opposition of Jupiter observed from Barbados. *J. Br. Astron. Assoc.* **121**, 337–340 (2011).
33. Cano, J. A. *LAlA: Laboratorio de Análisis de Imágenes Astronómicas* (Grup d'Estudis Astronòmics, Barcelona, 1998).
34. Bell, J. F. III et al. Mars surface mineralogy from Hubble Space Telescope imaging during 1994–1995: observations, calibration, and initial results. *J. Geophys. Res.* **102**, 9109–9123 (1997).
35. de Pater, I. et al. A multi-wavelength study of the 2009 impact on Jupiter: comparison of high-resolution images from Gemini, Keck and HST. *Icarus* **210**, 722–741 (2010).
36. Colina, L., Bohlin, R. C. & Castelli, F. The 0.12–2.5 micron absolute flux distribution of the Sun for comparison with solar analog stars. *Astron. J.* **112**, 307–315 (1996).
37. Mallama, A. The magnitude and albedo of Mars. *Icarus* **192**, 404–416 (2007).
38. García Muñoz, A. & Mills, F. P. Pre-conditioned backward Monte Carlo solutions to radiative transport in planetary atmospheres. Fundamentals: sampling of propagation directions in polarising media. *Astron. Astrophys.* **573**, A72 (2015).
39. Santer, R., Deschamps, M., Ksanfomaliti, L. V. & Dollfus, A. Photopolarimetry of Martian aerosols. II. Limb and terminator measurements. *Astron. Astrophys.* **158**, 247–258 (1986).
40. de Pater, I. & Lissauer, J. J. *Planetary Sciences* 2nd edn (Cambridge Univ. Press, 2010).



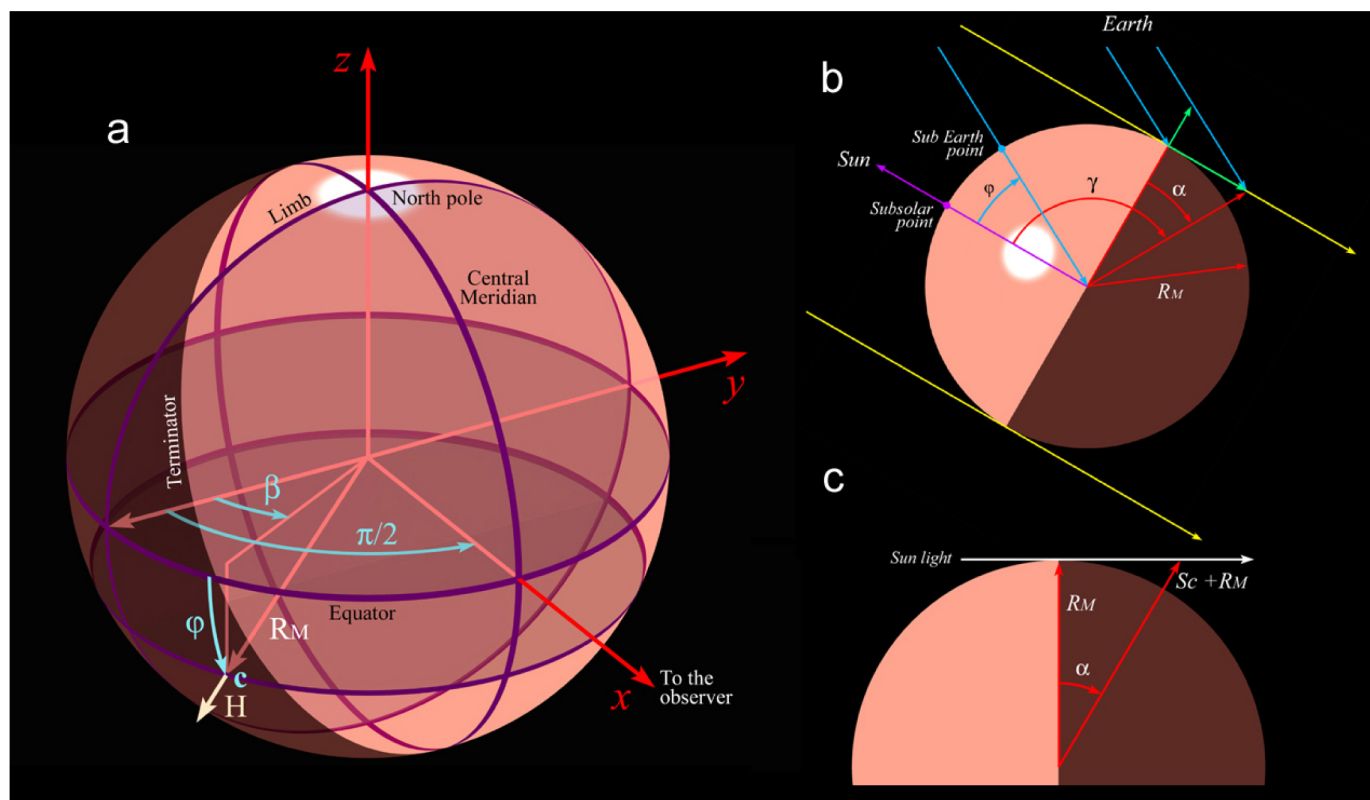
Extended Data Figure 1 | Images of the 2012 plume event (ringed) on 12–20 March. Dates in March (and authors) are as follows: **a**, 12 (M.D.); **b**, 15 (D. Peach); **c–f**, 20, plume in rotation (W.J.). Time indicated at top left of each panel is in UTC.



Extended Data Figure 2 | Images of the 2012 plume events (ringed) on 22 March and 13 April. **a**, First event on 22 March, 04:12 UTC (image by W.J.); **b**, second event on 13 April, 20:03 UTC (image by D. Peach).

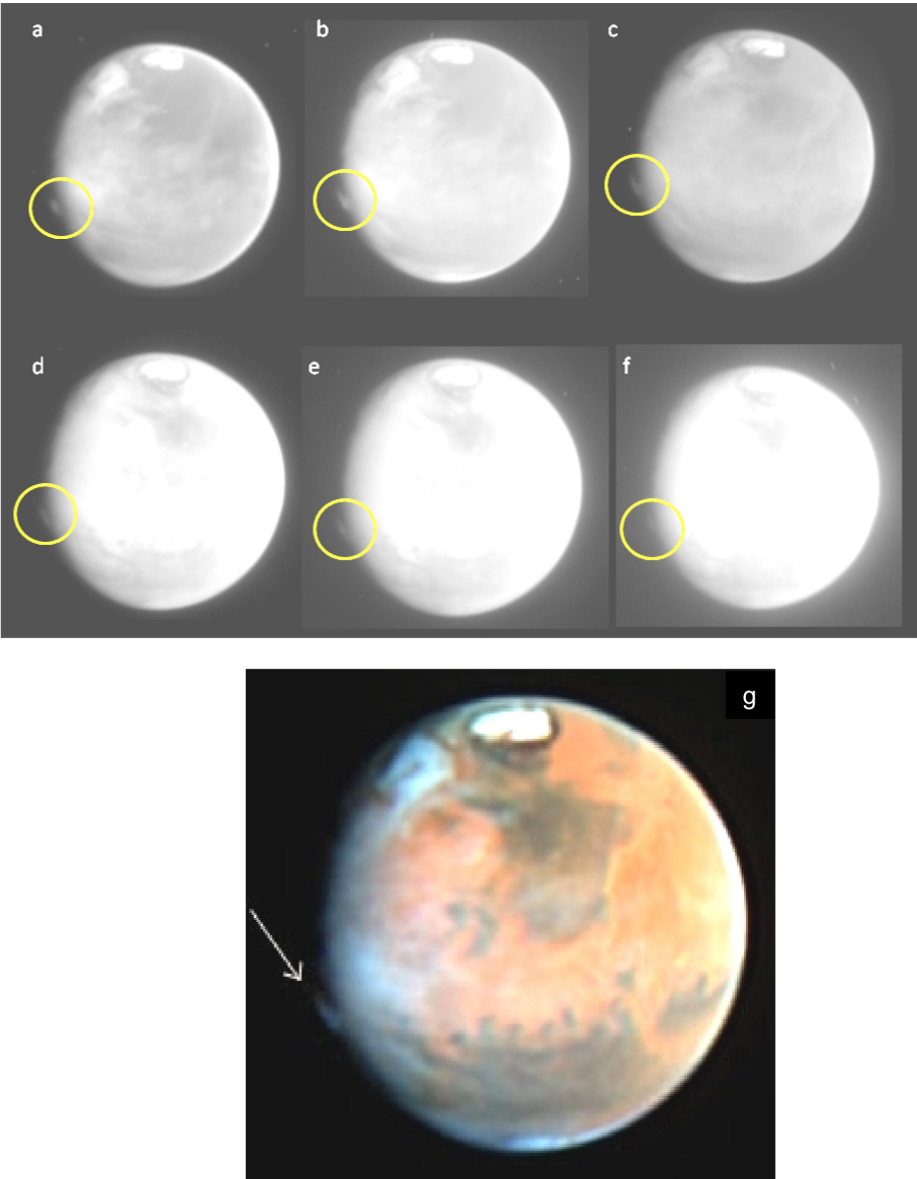


Extended Data Figure 3 | Images of the 2012 plume event (ringed) at different wavelengths on 21 March. a–c, Images by J.P., d–f, images by D. Parker, with filters indicated: a, d, B (blue); b, e, G (green); c, f, R (red). Time indicated is in UTC.



Extended Data Figure 4 | Martian viewing geometry. **a**, Angle definitions with the simulated protrusion of altitude H located at point c and out of the illuminated part of the disk near the limb. **b**, Top view, taking as a reference the planet's terminator and definition of γ and α angles when the cloud is on the equator (but the latitude of the subsolar point is not zero). Green arrows represent the projected cloud altitude as seen from Earth in the extreme

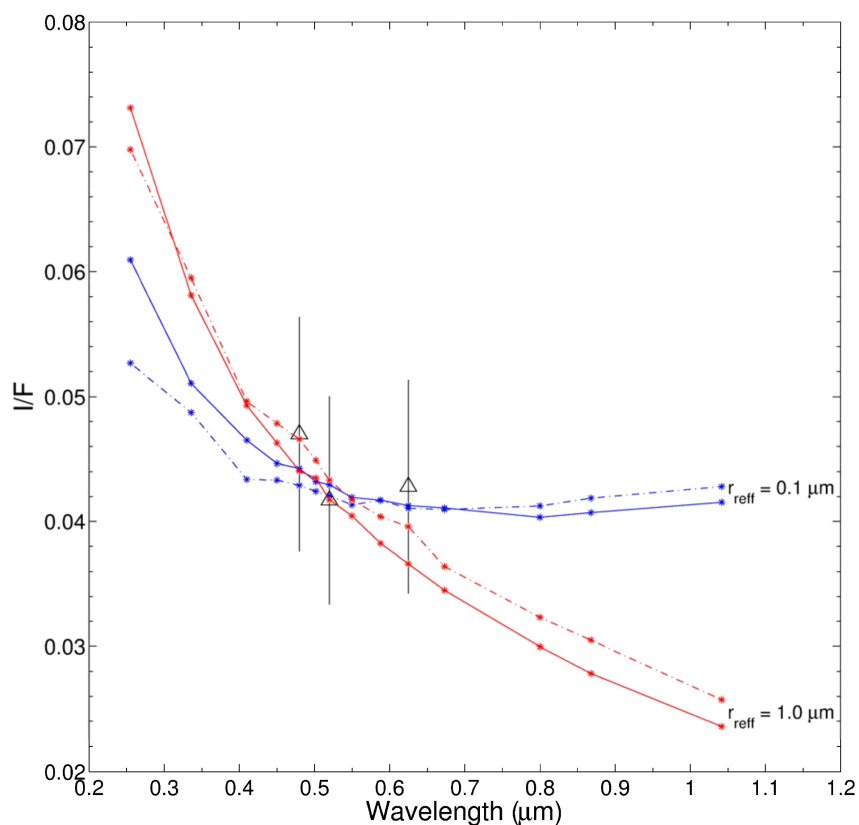
situations when the cloud is on the terminator and follows the grazing sunlight, and when it follows the planet's radius. To simplify the figure, and without loss of generality, the sub-Earth point is placed on the arc linking the subsolar point and the cloud base. **c**, General side view of the geometry of the planet's projected shadow.



Identification	HST number	Filter	Wavelength (nm)	FWHM (nm)
a	u3gi7701m.fits	F255W	257.3	42.6
b	u3gi7703m.fits	F410M	408.8	18.2
c	u3gi7704m.fits	F502N	5012.2	3.6
e	u3gi7705m.fits	F673N	673.2	6.3
f	u3gi7707m.fits	F1042M	1045.3	89.7
d	u3gi7708m.fits	F588N	589.3	6.4

Extended Data Figure 5 | Hubble Space Telescope images of the event on May 17 1997. Wavelengths and times in UTC were: **a**, 255 nm (17:27); **b**, 410 nm (17:35); **c**, 502 nm (17:38); **d**, 588 nm (17:50); **e**, 673 nm (17:41);

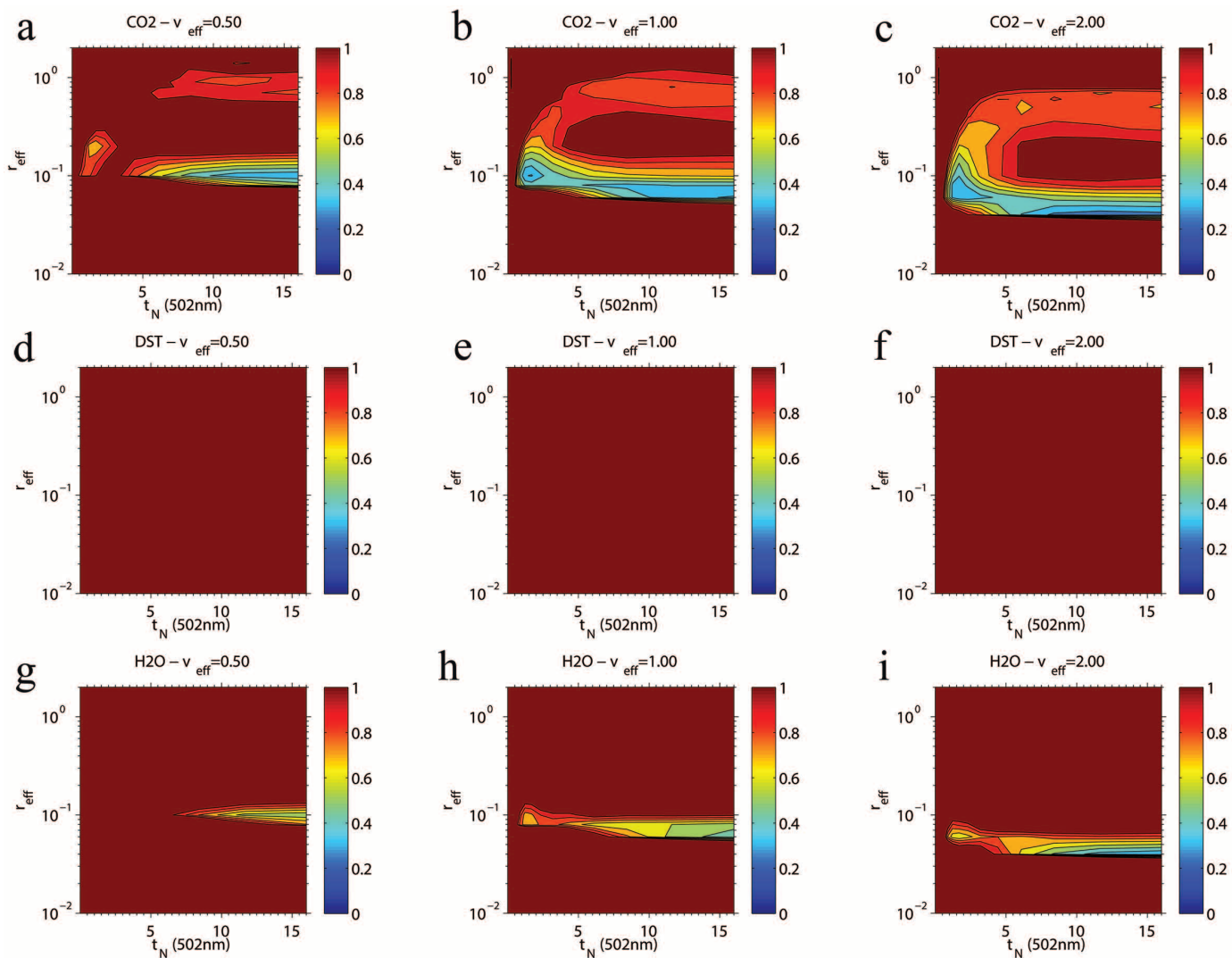
f, 1,042 nm (17:47); **g**, Colour composite. Plume ringed in **a–f**, arrowed in **g**. Table at bottom identifies each image and its HST number, and also shows filters used, giving their central wavelength and bandwidth (FWHM).



Extended Data Figure 6 | Radiative transfer model fit for the 2012 event.

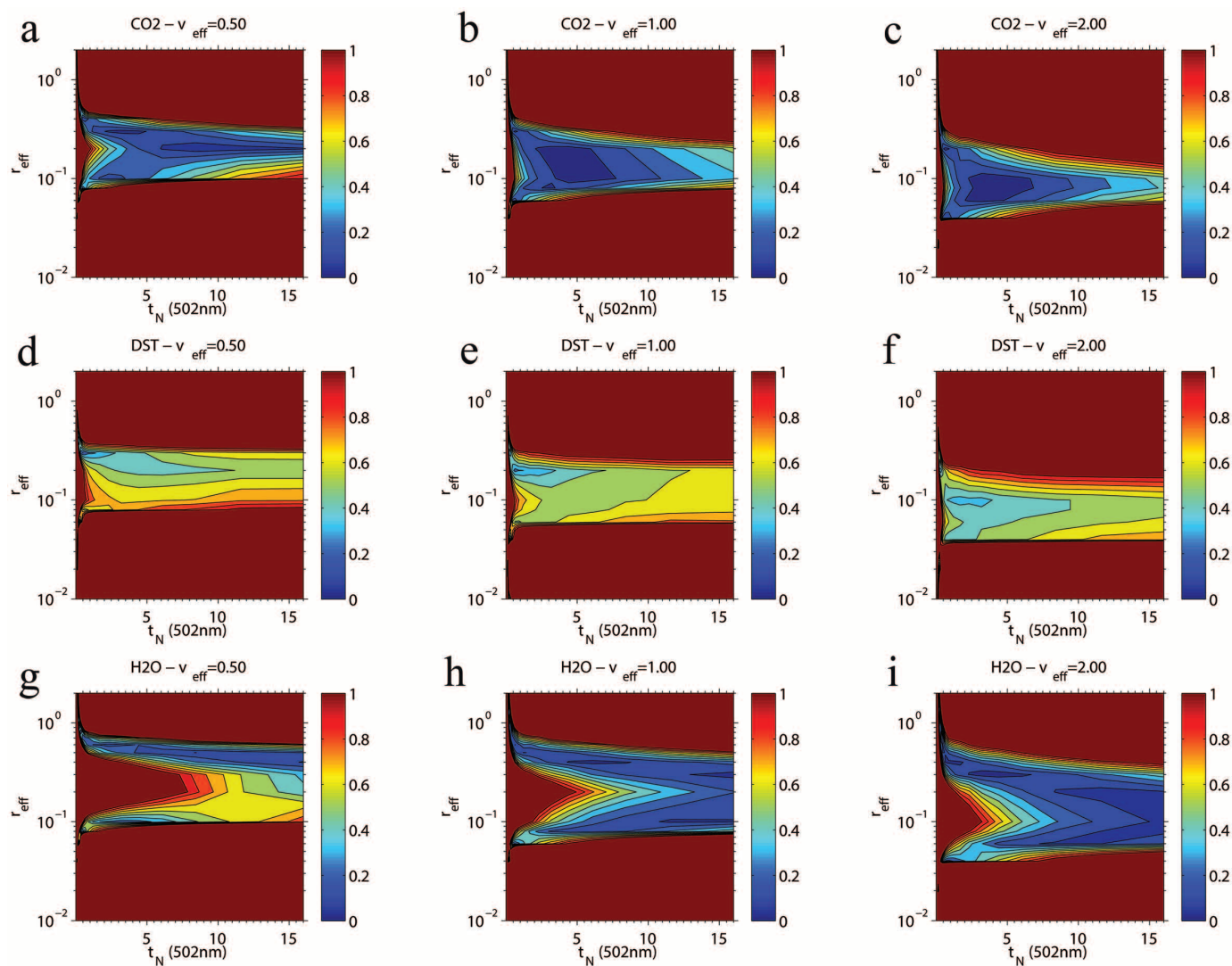
This is an example of the degeneracy of the model solution due to the narrow wavelength range covered in the 2012 event. Model fit as follows: solid lines, CO₂; dot-dashed lines, H₂O particles; blue, data for an effective radius

$r_{\text{eff}} = 0.1 \mu\text{m}$; red, data for $r_{\text{eff}} = 1.0 \mu\text{m}$; stars, wavelengths used in the calculations. As in Fig. 3, open black triangles show the observed reflectivity of the 2012 cloud. The error bars represent the average quadratic deviation of the measured reflectivity in the integration box.



Extended Data Figure 7 | Assessment of the radiative transfer model fit for the 1997 event. **a–i,** Colours show values of χ^2 (for measured I/F versus model calculation, colour scale at right) for the effective radius (r_{eff} in μm) versus optical depth (τ_N at 502 nm), and for different particle types and values of the indicated particle variance (v_{eff} shown in parentheses) as follows. **a,** CO_2 (0.5);

b, CO_2 (1.0); **c,** CO_2 (2.0); **d,** DST (dust, 0.5); **e,** DST (dust, 1.0); **f,** DST (dust, 2.0); **g,** H_2O (0.5); **h,** H_2O (1.0); and **i,** H_2O (2.0). The calculations are for a vertical extension of the cloud with $D = 100\text{ km}$, and they provide the best-fitting values of the whole free parameter space.



Extended Data Figure 8 | Assessment of the radiative transfer model fit for the 2012 event. As Extended Data Fig. 7 in terms of variables plotted, particle types and particle variance, but for the 2012 event.

Human-level control through deep reinforcement learning

Volodymyr Mnih^{1*}, Koray Kavukcuoglu^{1*}, David Silver^{1*}, Andrei A. Rusu¹, Joel Veness¹, Marc G. Bellemare¹, Alex Graves¹, Martin Riedmiller¹, Andreas K. Fiedjeland¹, Georg Ostrovski¹, Stig Petersen¹, Charles Beattie¹, Amir Sadik¹, Ioannis Antonoglou¹, Helen King¹, Dharshan Kumaran¹, Daan Wierstra¹, Shane Legg¹ & Demis Hassabis¹

The theory of reinforcement learning provides a normative account¹, deeply rooted in psychological² and neuroscientific³ perspectives on animal behaviour, of how agents may optimize their control of an environment. To use reinforcement learning successfully in situations approaching real-world complexity, however, agents are confronted with a difficult task: they must derive efficient representations of the environment from high-dimensional sensory inputs, and use these to generalize past experience to new situations. Remarkably, humans and other animals seem to solve this problem through a harmonious combination of reinforcement learning and hierarchical sensory processing systems^{4,5}, the former evidenced by a wealth of neural data revealing notable parallels between the phasic signals emitted by dopaminergic neurons and temporal difference reinforcement learning algorithms³. While reinforcement learning agents have achieved some successes in a variety of domains^{6–8}, their applicability has previously been limited to domains in which useful features can be handcrafted, or to domains with fully observed, low-dimensional state spaces. Here we use recent advances in training deep neural networks^{9–11} to develop a novel artificial agent, termed a deep Q-network, that can learn successful policies directly from high-dimensional sensory inputs using end-to-end reinforcement learning. We tested this agent on the challenging domain of classic Atari 2600 games¹². We demonstrate that the deep Q-network agent, receiving only the pixels and the game score as inputs, was able to surpass the performance of all previous algorithms and achieve a level comparable to that of a professional human games tester across a set of 49 games, using the same algorithm, network architecture and hyperparameters. This work bridges the divide between high-dimensional sensory inputs and actions, resulting in the first artificial agent that is capable of learning to excel at a diverse array of challenging tasks.

We set out to create a single algorithm that would be able to develop a wide range of competencies on a varied range of challenging tasks—a central goal of general artificial intelligence¹³ that has eluded previous efforts^{8,14,15}. To achieve this, we developed a novel agent, a deep Q-network (DQN), which is able to combine reinforcement learning with a class of artificial neural network¹⁶ known as deep neural networks. Notably, recent advances in deep neural networks^{9–11}, in which several layers of nodes are used to build up progressively more abstract representations of the data, have made it possible for artificial neural networks to learn concepts such as object categories directly from raw sensory data. We use one particularly successful architecture, the deep convolutional network¹⁷, which uses hierarchical layers of tiled convolutional filters to mimic the effects of receptive fields—inspired by Hubel and Wiesel's seminal work on feedforward processing in early visual cortex¹⁸—thereby exploiting the local spatial correlations present in images, and building in robustness to natural transformations such as changes of viewpoint or scale.

We consider tasks in which the agent interacts with an environment through a sequence of observations, actions and rewards. The goal of the

agent is to select actions in a fashion that maximizes cumulative future reward. More formally, we use a deep convolutional neural network to approximate the optimal action-value function

$$Q^*(s, a) = \max_{\pi} \mathbb{E}_{\pi} [r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots | s_t = s, a_t = a, \pi],$$

which is the maximum sum of rewards r_t discounted by γ at each time-step t , achievable by a behaviour policy $\pi = P(a|s)$, after making an observation (s) and taking an action (a) (see Methods)¹⁹.

Reinforcement learning is known to be unstable or even to diverge when a nonlinear function approximator such as a neural network is used to represent the action-value (also known as Q) function²⁰. This instability has several causes: the correlations present in the sequence of observations, the fact that small updates to Q may significantly change the policy and therefore change the data distribution, and the correlations between the action-values (Q) and the target values $r + \gamma \max_{a'} Q(s', a')$. We address these instabilities with a novel variant of Q-learning, which uses two key ideas. First, we used a biologically inspired mechanism termed experience replay^{21–23} that randomizes over the data, thereby removing correlations in the observation sequence and smoothing over changes in the data distribution (see below for details). Second, we used an iterative update that adjusts the action-values (Q) towards target values that are only periodically updated, thereby reducing correlations with the target.

While other stable methods exist for training neural networks in the reinforcement learning setting, such as neural fitted Q-iteration²⁴, these methods involve the repeated training of networks *de novo* on hundreds of iterations. Consequently, these methods, unlike our algorithm, are too inefficient to be used successfully with large neural networks. We parameterize an approximate value function $Q(s, a; \theta_i)$ using the deep convolutional neural network shown in Fig. 1, in which θ_i are the parameters (that is, weights) of the Q-network at iteration i . To perform experience replay we store the agent's experiences $e_t = (s_t, a_t, r_t, s_{t+1})$ at each time-step t in a data set $D_t = \{e_1, \dots, e_t\}$. During learning, we apply Q-learning updates, on samples (or minibatches) of experience $(s, a, r, s') \sim U(D)$, drawn uniformly at random from the pool of stored samples. The Q-learning update at iteration i uses the following loss function:

$$L_i(\theta_i) = \mathbb{E}_{(s, a, r, s') \sim U(D)} \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s, a; \theta_i) \right)^2 \right]$$

in which γ is the discount factor determining the agent's horizon, θ_i are the parameters of the Q-network at iteration i and θ_i^- are the network parameters used to compute the target at iteration i . The target network parameters θ_i^- are only updated with the Q-network parameters (θ_i) every C steps and are held fixed between individual updates (see Methods).

To evaluate our DQN agent, we took advantage of the Atari 2600 platform, which offers a diverse array of tasks ($n = 49$) designed to be

¹Google DeepMind, 5 New Street Square, London EC4A 3TW, UK.

*These authors contributed equally to this work.

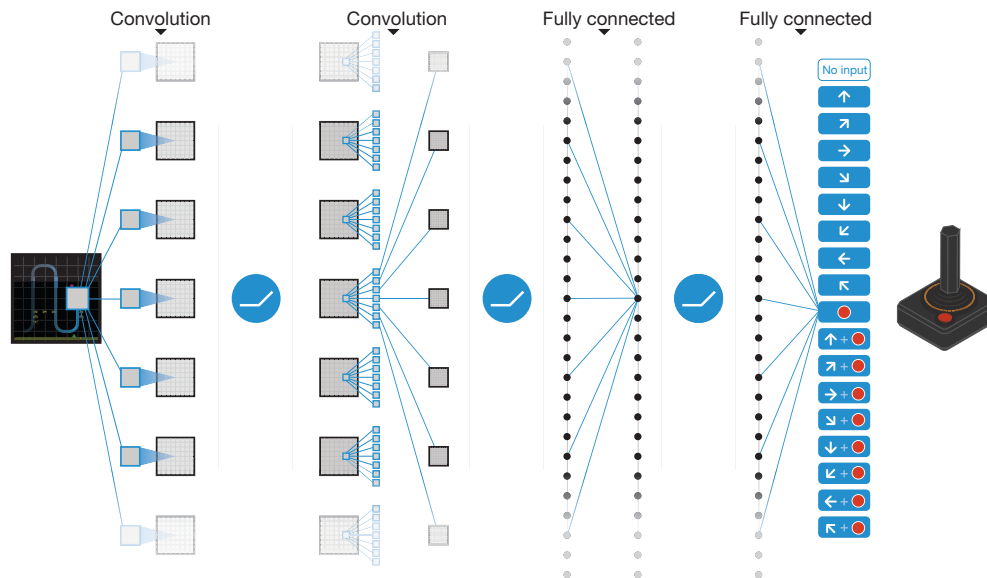


Figure 1 | Schematic illustration of the convolutional neural network. The details of the architecture are explained in the Methods. The input to the neural network consists of an $84 \times 84 \times 4$ image produced by the preprocessing map ϕ , followed by three convolutional layers (note: snaking blue line

symbolizes sliding of each filter across input image) and two fully connected layers with a single output for each valid action. Each hidden layer is followed by a rectifier nonlinearity (that is, $\max(0, x)$).

difficult and engaging for human players. We used the same network architecture, hyperparameter values (see Extended Data Table 1) and learning procedure throughout—taking high-dimensional data (210×160 colour video at 60 Hz) as input—to demonstrate that our approach robustly learns successful policies over a variety of games based solely on sensory inputs with only very minimal prior knowledge (that is, merely the input data were visual images, and the number of actions available in each game, but not their correspondences; see Methods). Notably, our method was able to train large neural networks using a reinforcement learning signal and stochastic gradient descent in a stable manner—illustrated by the temporal evolution of two indices of learning (the agent’s average score-per-episode and average predicted Q-values; see Fig. 2 and Supplementary Discussion for details).

We compared DQN with the best performing methods from the reinforcement learning literature on the 49 games where results were available^{12,15}. In addition to the learned agents, we also report scores for a professional human games tester playing under controlled conditions and a policy that selects actions uniformly at random (Extended Data Table 2 and Fig. 3, denoted by 100% (human) and 0% (random) on y axis; see Methods). Our DQN method outperforms the best existing reinforcement learning methods on 43 of the games without incorporating any of the additional prior knowledge about Atari 2600 games used by other approaches (for example, refs 12, 15). Furthermore, our DQN agent performed at a level that was comparable to that of a professional human games tester across the set of 49 games, achieving more than 75% of the human score on more than half of the games (29 games;

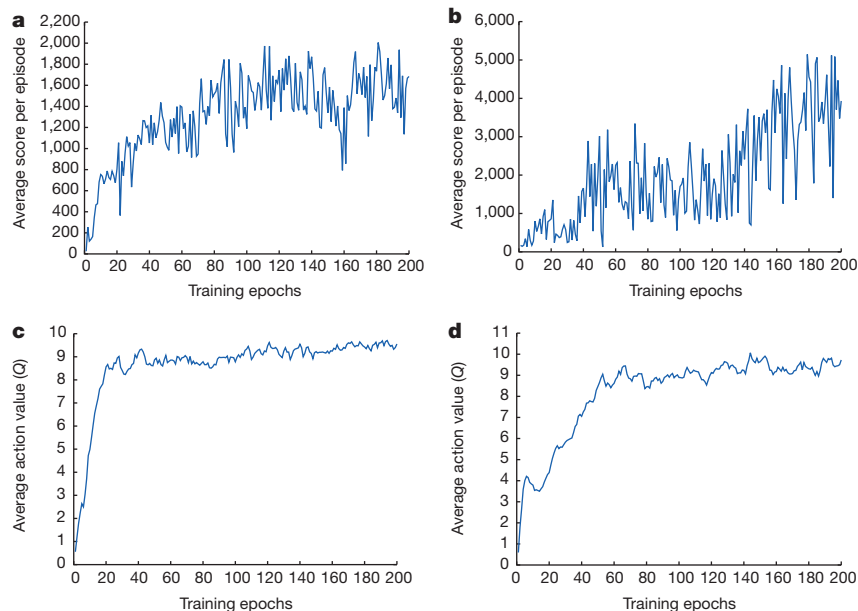


Figure 2 | Training curves tracking the agent’s average score and average predicted action-value. **a**, Each point is the average score achieved per episode after the agent is run with ϵ -greedy policy ($\epsilon = 0.05$) for 520 k frames on Space Invaders. **b**, Average score achieved per episode for Seaquest. **c**, Average predicted action-value on a held-out set of states on Space Invaders. Each point

on the curve is the average of the action-value Q computed over the held-out set of states. Note that Q -values are scaled due to clipping of rewards (see Methods). **d**, Average predicted action-value on Seaquest. See Supplementary Discussion for details.

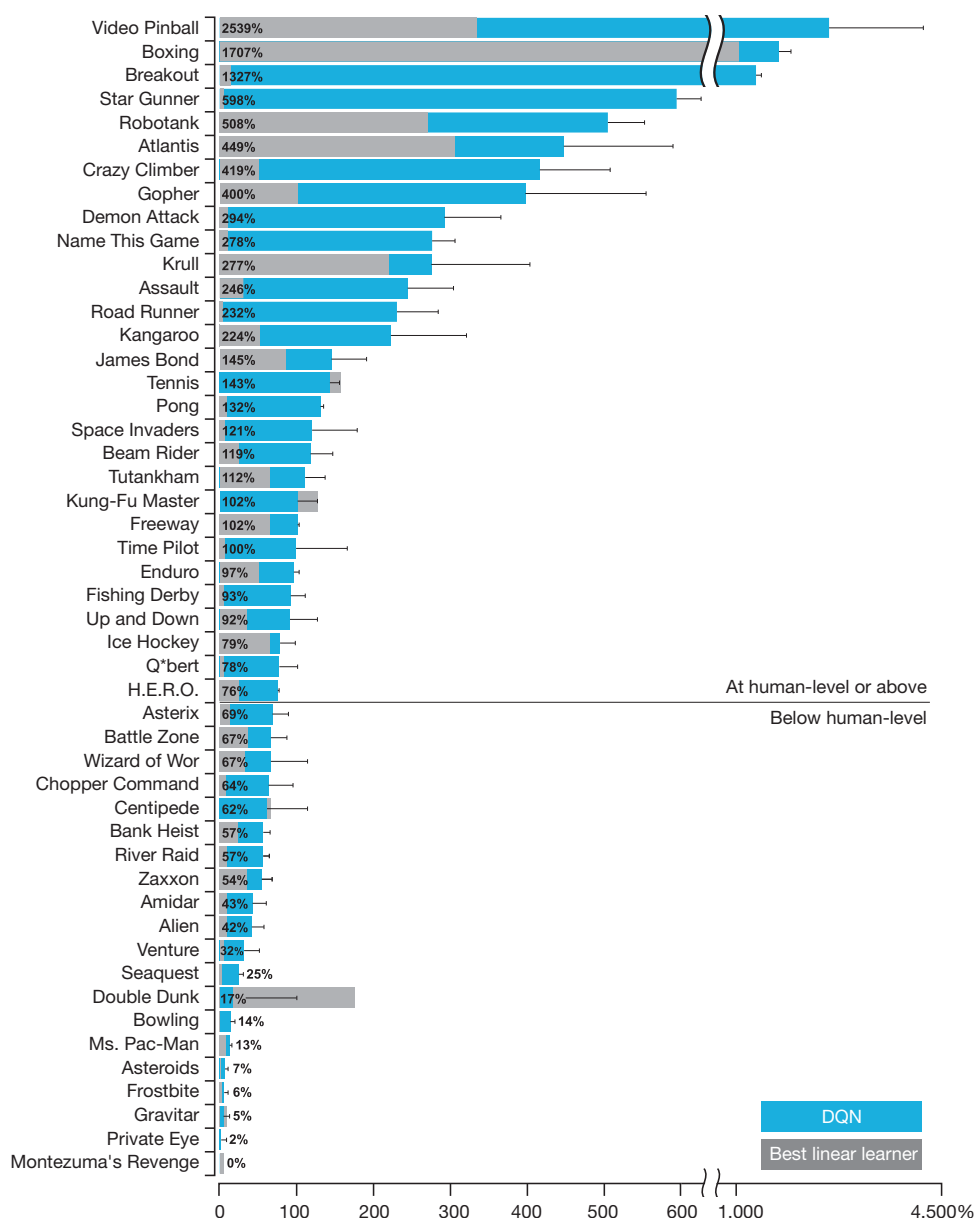


Figure 3 | Comparison of the DQN agent with the best reinforcement learning methods¹⁵ in the literature. The performance of DQN is normalized with respect to a professional human games tester (that is, 100% level) and random play (that is, 0% level). Note that the normalized performance of DQN, expressed as a percentage, is calculated as: $100 \times (\text{DQN score} - \text{random play score}) / (\text{human score} - \text{random play score})$. It can be seen that DQN

outperforms competing methods (also see Extended Data Table 2) in almost all the games, and performs at a level that is broadly comparable with or superior to a professional human games tester (that is, operationalized as a level of 75% or above) in the majority of games. Audio output was disabled for both human players and agents. Error bars indicate s.d. across the 30 evaluation episodes, starting with different initial conditions.

see Fig. 3, Supplementary Discussion and Extended Data Table 2). In additional simulations (see Supplementary Discussion and Extended Data Tables 3 and 4), we demonstrate the importance of the individual core components of the DQN agent—the replay memory, separate target Q-network and deep convolutional network architecture—by disabling them and demonstrating the detrimental effects on performance.

We next examined the representations learned by DQN that underpinned the successful performance of the agent in the context of the game Space Invaders (see Supplementary Video 1 for a demonstration of the performance of DQN), by using a technique developed for the visualization of high-dimensional data called 't-SNE'²²⁵ (Fig. 4). As expected, the t-SNE algorithm tends to map the DQN representation of perceptually similar states to nearby points. Interestingly, we also found instances in which the t-SNE algorithm generated similar embeddings for DQN representations of states that are close in terms of expected reward but

perceptually dissimilar (Fig. 4, bottom right, top left and middle), consistent with the notion that the network is able to learn representations that support adaptive behaviour from high-dimensional sensory inputs. Furthermore, we also show that the representations learned by DQN are able to generalize to data generated from policies other than its own—in simulations where we presented as input to the network game states experienced during human and agent play, recorded the representations of the last hidden layer, and visualized the embeddings generated by the t-SNE algorithm (Extended Data Fig. 1 and Supplementary Discussion). Extended Data Fig. 2 provides an additional illustration of how the representations learned by DQN allow it to accurately predict state and action values.

It is worth noting that the games in which DQN excels are extremely varied in their nature, from side-scrolling shooters (River Raid) to boxing games (Boxing) and three-dimensional car-racing games (Enduro).

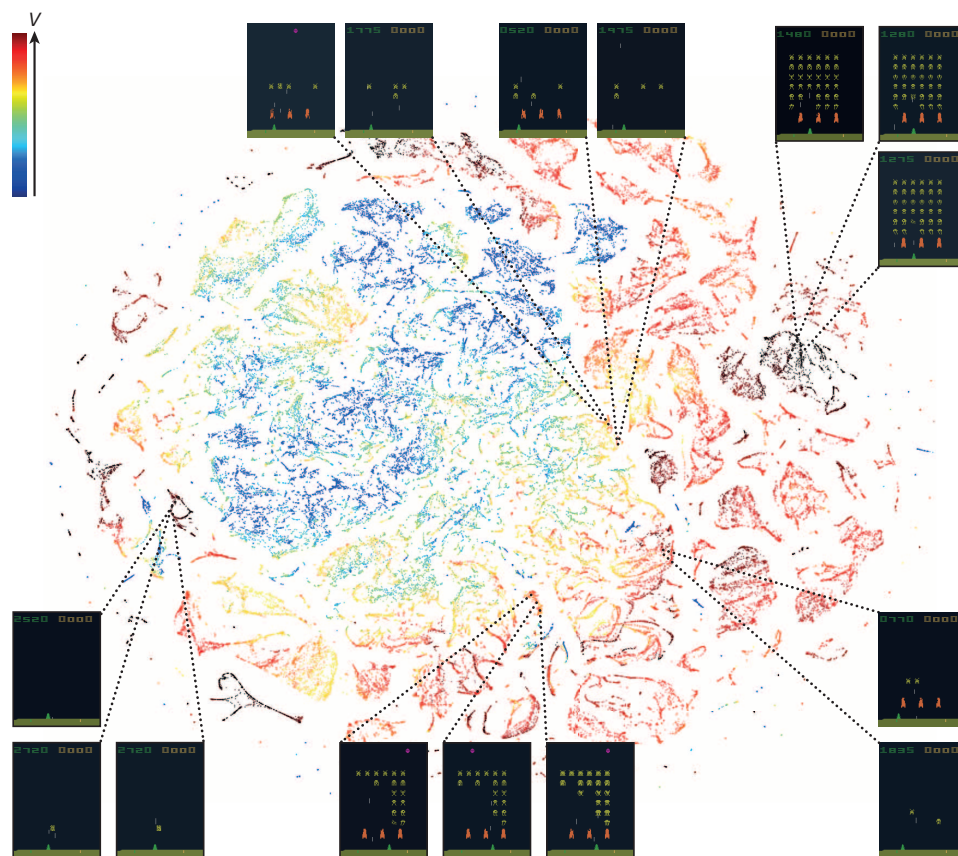


Figure 4 | Two-dimensional t-SNE embedding of the representations in the last hidden layer assigned by DQN to game states experienced while playing Space Invaders. The plot was generated by letting the DQN agent play for 2 h of real game time and running the t-SNE algorithm²⁵ on the last hidden layer representations assigned by DQN to each experienced game state. The points are coloured according to the state values (V , maximum expected reward of a state) predicted by DQN for the corresponding game states (ranging from dark red (highest V) to dark blue (lowest V)). The screenshots corresponding to a selected number of points are shown. The DQN agent

predicts high state values for both full (top right screenshots) and nearly complete screens (bottom left screenshots) because it has learned that completing a screen leads to a new screen full of enemy ships. Partially completed screens (bottom screenshots) are assigned lower state values because less immediate reward is available. The screens shown on the bottom right and top left and middle are less perceptually similar than the other examples but are still mapped to nearby representations and similar values because the orange bunkers do not carry great significance near the end of a level. With permission from Square Enix Limited.

Indeed, in certain games DQN is able to discover a relatively long-term strategy (for example, Breakout: the agent learns the optimal strategy, which is to first dig a tunnel around the side of the wall allowing the ball to be sent around the back to destroy a large number of blocks; see Supplementary Video 2 for illustration of development of DQN's performance over the course of training). Nevertheless, games demanding more temporally extended planning strategies still constitute a major challenge for all existing agents including DQN (for example, Montezuma's Revenge).

In this work, we demonstrate that a single architecture can successfully learn control policies in a range of different environments with only very minimal prior knowledge, receiving only the pixels and the game score as inputs, and using the same algorithm, network architecture and hyperparameters on each game, privy only to the inputs a human player would have. In contrast to previous work^{24,26}, our approach incorporates 'end-to-end' reinforcement learning that uses reward to continuously shape representations within the convolutional network towards salient features of the environment that facilitate value estimation. This principle draws on neurobiological evidence that reward signals during perceptual learning may influence the characteristics of representations within primate visual cortex^{27,28}. Notably, the successful integration of reinforcement learning with deep network architectures was critically dependent on our incorporation of a replay algorithm^{21–23} involving the storage and representation of recently experienced transitions. Convergent evidence suggests that the hippocampus may support the physical

realization of such a process in the mammalian brain, with the time-compressed reactivation of recently experienced trajectories during offline periods^{21,22} (for example, waking rest) providing a putative mechanism by which value functions may be efficiently updated through interactions with the basal ganglia²². In the future, it will be important to explore the potential use of biasing the content of experience replay towards salient events, a phenomenon that characterizes empirically observed hippocampal replay²⁹, and relates to the notion of 'prioritized sweeping'³⁰ in reinforcement learning. Taken together, our work illustrates the power of harnessing state-of-the-art machine learning techniques with biologically inspired mechanisms to create agents that are capable of learning to master a diverse array of challenging tasks.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 July 2014; accepted 16 January 2015.

1. Sutton, R. & Barto, A. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
2. Thorndike, E. L. *Animal Intelligence: Experimental studies* (Macmillan, 1911).
3. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
4. Serre, T., Wolf, L. & Poggio, T. Object recognition with features inspired by visual cortex. *Proc. IEEE. Comput. Soc. Conf. Comput. Vis. Pattern. Recognit.* 994–1000 (2005).
5. Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202 (1980).

6. Tesauro, G. Temporal difference learning and TD-Gammon. *Commun. ACM* **38**, 58–68 (1995).
7. Riedmiller, M., Gabel, T., Hafner, R. & Lange, S. Reinforcement learning for robot soccer. *Auton. Robots* **27**, 55–73 (2009).
8. Diuk, C., Cohen, A. & Littman, M. L. An object-oriented representation for efficient reinforcement learning. *Proc. Int. Conf. Mach. Learn.* 240–247 (2008).
9. Bengio, Y. Learning deep architectures for AI. *Foundations and Trends in Machine Learning* **2**, 1–127 (2009).
10. Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**, 1106–1114 (2012).
11. Hinton, G. E. & Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **313**, 504–507 (2006).
12. Bellemare, M. G., Naddaf, Y., Veness, J. & Bowling, M. The arcade learning environment: An evaluation platform for general agents. *J. Artif. Intell. Res.* **47**, 253–279 (2013).
13. Legg, S. & Hutter, M. Universal Intelligence: a definition of machine intelligence. *Minds Mach.* **17**, 391–444 (2007).
14. Genesereth, M., Love, N. & Pell, B. General game playing: overview of the AAAI competition. *AI Mag.* **26**, 62–72 (2005).
15. Bellemare, M. G., Veness, J. & Bowling, M. Investigating contingency awareness using Atari 2600 games. *Proc. Conf. AAAI. Artif. Intell.* 864–871 (2012).
16. McClelland, J. L., Rumelhart, D. E. & Group, T. P. R. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, 1986).
17. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
18. Hubel, D. H. & Wiesel, T. N. Shape and arrangement of columns in cat's striate cortex. *J. Physiol.* **165**, 559–568 (1963).
19. Watkins, C. J. & Dayan, P. Q-learning. *Mach. Learn.* **8**, 279–292 (1992).
20. Tsitsiklis, J. & Roy, B. V. An analysis of temporal-difference learning with function approximation. *IEEE Trans. Automat. Contr.* **42**, 674–690 (1997).
21. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
22. O'Neill, J., Pleydell-Bouverie, B., Dupret, D. & Csicsvari, J. Play it again: reactivation of waking experience and memory. *Trends Neurosci.* **33**, 220–229 (2010).
23. Lin, L.-J. Reinforcement learning for robots using neural networks. Technical Report, DTIC Document (1993).
24. Riedmiller, M. Neural fitted Q iteration - first experiences with a data efficient neural reinforcement learning method. *Mach. Learn.: ECML*, **3720**, 317–328 (Springer, 2005).
25. Van der Maaten, L. J. P. & Hinton, G. E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
26. Lange, S. & Riedmiller, M. Deep auto-encoder neural networks in reinforcement learning. *Proc. Int. Jt. Conf. Neural. Netw.* 1–8 (2010).
27. Law, C.-T. & Gold, J. I. Reinforcement learning can account for associative and perceptual learning on a visual decision task. *Nature Neurosci.* **12**, 655 (2009).
28. Sigala, N. & Logothetis, N. K. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* **415**, 318–320 (2002).
29. Bendor, D. & Wilson, M. A. Biasing the content of hippocampal replay during sleep. *Nature Neurosci.* **15**, 1439–1444 (2012).
30. Moore, A. & Atkeson, C. Prioritized sweeping: reinforcement learning with less data and less real time. *Mach. Learn.* **13**, 103–130 (1993).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Hinton, P. Dayan and M. Bowling for discussions, A. Cain and J. Keene for work on the visuals, K. Keller and P. Rogers for help with the visuals, G. Wayne for comments on an earlier version of the manuscript, and the rest of the DeepMind team for their support, ideas and encouragement.

Author Contributions V.M., K.K., D.S., J.V., M.G.B., M.R., A.G., D.W., S.L. and D.H. conceptualized the problem and the technical framework. V.M., K.K., A.A.R. and D.S. developed and tested the algorithms. J.V., S.P., C.B., A.A.R., M.G.B., I.A., A.K.F., G.O. and A.S. created the testing platform. K.K., H.K., S.L. and D.H. managed the project. K.K., D.K., D.H., V.M., D.S., A.G., A.A.R., J.V. and M.G.B. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.K. (korayk@google.com) or D.H. (demishassabis@google.com).

METHODS

Preprocessing. Working directly with raw Atari 2600 frames, which are 210×160 pixel images with a 128-colour palette, can be demanding in terms of computation and memory requirements. We apply a basic preprocessing step aimed at reducing the input dimensionality and dealing with some artefacts of the Atari 2600 emulator. First, to encode a single frame we take the maximum value for each pixel colour value over the frame being encoded and the previous frame. This was necessary to remove flickering that is present in games where some objects appear only in even frames while other objects appear only in odd frames, an artefact caused by the limited number of sprites Atari 2600 can display at once. Second, we then extract the Y channel, also known as luminance, from the RGB frame and rescale it to 84×84 . The function ϕ from algorithm 1 described below applies this preprocessing to the m most recent frames and stacks them to produce the input to the Q-function, in which $m = 4$, although the algorithm is robust to different values of m (for example, 3 or 5).

Code availability. The source code can be accessed at <https://sites.google.com/a/deepmind.com/dqn> for non-commercial uses only.

Model architecture. There are several possible ways of parameterizing Q using a neural network. Because Q maps history-action pairs to scalar estimates of their Q-value, the history and the action have been used as inputs to the neural network by some previous approaches^{24,26}. The main drawback of this type of architecture is that a separate forward pass is required to compute the Q-value of each action, resulting in a cost that scales linearly with the number of actions. We instead use an architecture in which there is a separate output unit for each possible action, and only the state representation is an input to the neural network. The outputs correspond to the predicted Q-values of the individual actions for the input state. The main advantage of this type of architecture is the ability to compute Q-values for all possible actions in a given state with only a single forward pass through the network.

The exact architecture, shown schematically in Fig. 1, is as follows. The input to the neural network consists of an $84 \times 84 \times 4$ image produced by the preprocessing map ϕ . The first hidden layer convolves 32 filters of 8×8 with stride 4 with the input image and applies a rectifier nonlinearity^{31,32}. The second hidden layer convolves 64 filters of 4×4 with stride 2, again followed by a rectifier nonlinearity. This is followed by a third convolutional layer that convolves 64 filters of 3×3 with stride 1 followed by a rectifier. The final hidden layer is fully-connected and consists of 512 rectifier units. The output layer is a fully-connected linear layer with a single output for each valid action. The number of valid actions varied between 4 and 18 on the games we considered.

Training details. We performed experiments on 49 Atari 2600 games where results were available for all other comparable methods^{12,15}. A different network was trained on each game: the same network architecture, learning algorithm and hyperparameter settings (see Extended Data Table 1) were used across all games, showing that our approach is robust enough to work on a variety of games while incorporating only minimal prior knowledge (see below). While we evaluated our agents on unmodified games, we made one change to the reward structure of the games during training only. As the scale of scores varies greatly from game to game, we clipped all positive rewards at 1 and all negative rewards at -1 , leaving 0 rewards unchanged. Clipping the rewards in this manner limits the scale of the error derivatives and makes it easier to use the same learning rate across multiple games. At the same time, it could affect the performance of our agent since it cannot differentiate between rewards of different magnitude. For games where there is a life counter, the Atari 2600 emulator also sends the number of lives left in the game, which is then used to mark the end of an episode during training.

In these experiments, we used the RMSProp (see http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf) algorithm with minibatches of size 32. The behaviour policy during training was ϵ -greedy with ϵ annealed linearly from 1.0 to 0.1 over the first million frames, and fixed at 0.1 thereafter. We trained for a total of 50 million frames (that is, around 38 days of game experience in total) and used a replay memory of 1 million most recent frames.

Following previous approaches to playing Atari 2600 games, we also use a simple frame-skipping technique¹⁵. More precisely, the agent sees and selects actions on every k th frame instead of every frame, and its last action is repeated on skipped frames. Because running the emulator forward for one step requires much less computation than having the agent select an action, this technique allows the agent to play roughly k times more games without significantly increasing the runtime. We use $k = 4$ for all games.

The values of all the hyperparameters and optimization parameters were selected by performing an informal search on the games Pong, Breakout, Seaquest, Space Invaders and Beam Rider. We did not perform a systematic grid search owing to the high computational cost. These parameters were then held fixed across all other games. The values and descriptions of all hyperparameters are provided in Extended Data Table 1.

Our experimental setup amounts to using the following minimal prior knowledge: that the input data consisted of visual images (motivating our use of a convolutional deep network), the game-specific score (with no modification), number of actions, although not their correspondences (for example, specification of the up 'button') and the life count.

Evaluation procedure. The trained agents were evaluated by playing each game 30 times for up to 5 min each time with different initial random conditions ('no-op'; see Extended Data Table 1) and an ϵ -greedy policy with $\epsilon = 0.05$. This procedure is adopted to minimize the possibility of overfitting during evaluation. The random agent served as a baseline comparison and chose a random action at 10 Hz which is every sixth frame, repeating its last action on intervening frames. 10 Hz is about the fastest that a human player can select the 'fire' button, and setting the random agent to this frequency avoids spurious baseline scores in a handful of the games. We did also assess the performance of a random agent that selected an action at 60 Hz (that is, every frame). This had a minimal effect: changing the normalized DQN performance by more than 5% in only six games (Boxing, Breakout, Crazy Climber, Demon Attack, Krull and Robotank), and in all these games DQN outperformed the expert human by a considerable margin.

The professional human tester used the same emulator engine as the agents, and played under controlled conditions. The human tester was not allowed to pause, save or reload games. As in the original Atari 2600 environment, the emulator was run at 60 Hz and the audio output was disabled: as such, the sensory input was equated between human player and agents. The human performance is the average reward achieved from around 20 episodes of each game lasting a maximum of 5 min each, following around 2 h of practice playing each game.

Algorithm. We consider tasks in which an agent interacts with an environment, in this case the Atari emulator, in a sequence of actions, observations and rewards. At each time-step the agent selects an action a_t from the set of legal game actions, $\mathcal{A} = \{1, \dots, K\}$. The action is passed to the emulator and modifies its internal state and the game score. In general the environment may be stochastic. The emulator's internal state is not observed by the agent; instead the agent observes an image $x_t \in \mathbb{R}^d$ from the emulator, which is a vector of pixel values representing the current screen. In addition it receives a reward r_t representing the change in game score. Note that in general the game score may depend on the whole previous sequence of actions and observations; feedback about an action may only be received after many thousands of time-steps have elapsed.

Because the agent only observes the current screen, the task is partially observed³³ and many emulator states are perceptually aliased (that is, it is impossible to fully understand the current situation from only the current screen x_t). Therefore, sequences of actions and observations, $s_t = x_1, a_1, x_2, \dots, a_{t-1}, x_t$, are input to the algorithm, which then learns game strategies depending upon these sequences. All sequences in the emulator are assumed to terminate in a finite number of time-steps. This formalism gives rise to a large but finite Markov decision process (MDP) in which each sequence is a distinct state. As a result, we can apply standard reinforcement learning methods for MDPs, simply by using the complete sequence s_t as the state representation at time t .

The goal of the agent is to interact with the emulator by selecting actions in a way that maximizes future rewards. We make the standard assumption that future rewards are discounted by a factor of γ per time-step (γ was set to 0.99 throughout), and

define the future discounted return at time t as $R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'}$, in which T is the time-step at which the game terminates. We define the optimal action-value function $Q^*(s, a)$ as the maximum expected return achievable by following any policy, after seeing some sequence s and then taking some action a , $Q^*(s, a) = \max_{\pi} \mathbb{E}[R_t | s_t = s, a_t = a, \pi]$ in which π is a policy mapping sequences to actions (or distributions over actions).

The optimal action-value function obeys an important identity known as the Bellman equation. This is based on the following intuition: if the optimal value $Q^*(s', a')$ of the sequence s' at the next time-step was known for all possible actions a' , then the optimal strategy is to select the action a' maximizing the expected value of $r + \gamma Q^*(s', a')$:

$$Q^*(s, a) = \mathbb{E}_s \left[r + \gamma \max_{a'} Q^*(s', a') | s, a \right]$$

The basic idea behind many reinforcement learning algorithms is to estimate the action-value function by using the Bellman equation as an iterative update, $Q_{i+1}(s, a) = \mathbb{E}_s [r + \gamma \max_{a'} Q_i(s', a') | s, a]$. Such value iteration algorithms converge to the optimal action-value function, $Q_i \rightarrow Q^*$ as $i \rightarrow \infty$. In practice, this basic approach is impractical, because the action-value function is estimated separately for each sequence, without any generalization. Instead, it is common to use a function approximator to estimate the action-value function, $Q(s, a; \theta) \approx Q^*(s, a)$. In the reinforcement learning community this is typically a linear function approximator, but

sometimes a nonlinear function approximator is used instead, such as a neural network. We refer to a neural network function approximator with weights θ as a Q-network. A Q-network can be trained by adjusting the parameters θ_i at iteration i to reduce the mean-squared error in the Bellman equation, where the optimal target values $r + \gamma \max_{a'} Q^*(s', a')$ are substituted with approximate target values $y = r + \gamma \max_{a'} Q(s', a'; \theta_i^-)$, using parameters θ_i^- from some previous iteration. This leads to a sequence of loss functions $L_i(\theta_i)$ that changes at each iteration i ,

$$\begin{aligned} L_i(\theta_i) &= \mathbb{E}_{s,a,r} [(\mathbb{E}_{s'} [y|s,a] - Q(s,a; \theta_i))^2] \\ &= \mathbb{E}_{s,a,r,s'} [(y - Q(s,a; \theta_i))^2] + \mathbb{E}_{s,a,r} [\mathbb{V}_{s'} [y]]. \end{aligned}$$

Note that the targets depend on the network weights; this is in contrast with the targets used for supervised learning, which are fixed before learning begins. At each stage of optimization, we hold the parameters from the previous iteration θ_i^- fixed when optimizing the i th loss function $L_i(\theta_i)$, resulting in a sequence of well-defined optimization problems. The final term is the variance of the targets, which does not depend on the parameters θ_i that we are currently optimizing, and may therefore be ignored. Differentiating the loss function with respect to the weights we arrive at the following gradient:

$$\nabla_{\theta_i} L(\theta_i) = \mathbb{E}_{s,a,r,s'} \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s,a; \theta_i) \right) \nabla_{\theta_i} Q(s,a; \theta_i) \right].$$

Rather than computing the full expectations in the above gradient, it is often computationally expedient to optimize the loss function by stochastic gradient descent. The familiar Q-learning algorithm¹⁹ can be recovered in this framework by updating the weights after every time step, replacing the expectations using single samples, and setting $\theta_i^- = \theta_{i-1}$.

Note that this algorithm is model-free: it solves the reinforcement learning task directly using samples from the emulator, without explicitly estimating the reward and transition dynamics $P(r, s' | s, a)$. It is also off-policy: it learns about the greedy policy $a = \arg\max_{a'} Q(s, a'; \theta)$, while following a behaviour distribution that ensures adequate exploration of the state space. In practice, the behaviour distribution is often selected by an ε -greedy policy that follows the greedy policy with probability $1 - \varepsilon$ and selects a random action with probability ε .

Training algorithm for deep Q-networks. The full algorithm for training deep Q-networks is presented in Algorithm 1. The agent selects and executes actions according to an ε -greedy policy based on Q . Because using histories of arbitrary length as inputs to a neural network can be difficult, our Q-function instead works on a fixed length representation of histories produced by the function ϕ described above. The algorithm modifies standard online Q-learning in two ways to make it suitable for training large neural networks without diverging.

First, we use a technique known as experience replay²³ in which we store the agent's experiences at each time-step, $e_t = (s_t, a_t, r_t, s_{t+1})$, in a data set $D_t = \{e_1, \dots, e_t\}$, pooled over many episodes (where the end of an episode occurs when a terminal state is reached) into a replay memory. During the inner loop of the algorithm, we apply Q-learning updates, or minibatch updates, to samples of experience, $(s, a, r, s') \sim U(D)$, drawn at random from the pool of stored samples. This approach has several advantages over standard online Q-learning. First, each step of experience is potentially used in many weight updates, which allows for greater data efficiency. Second, learning directly from consecutive samples is inefficient, owing to the strong correlations between the samples; randomizing the samples breaks these correlations and therefore reduces the variance of the updates. Third, when learning on-policy the current parameters determine the next data sample that the parameters are trained on. For example, if the maximizing action is to move left then the training samples will be dominated by samples from the left-hand side; if the maximizing action then switches to the right then the training distribution will also switch. It is easy to see how unwanted feedback loops may arise and the parameters could get stuck in a poor local minimum, or even diverge catastrophically²⁰. By using experience

replay the behaviour distribution is averaged over many of its previous states, smoothing out learning and avoiding oscillations or divergence in the parameters. Note that when learning by experience replay, it is necessary to learn off-policy (because our current parameters are different to those used to generate the sample), which motivates the choice of Q-learning.

In practice, our algorithm only stores the last N experience tuples in the replay memory, and samples uniformly at random from D when performing updates. This approach is in some respects limited because the memory buffer does not differentiate important transitions and always overwrites with recent transitions owing to the finite memory size N . Similarly, the uniform sampling gives equal importance to all transitions in the replay memory. A more sophisticated sampling strategy might emphasize transitions from which we can learn the most, similar to prioritized sweeping³⁰.

The second modification to online Q-learning aimed at further improving the stability of our method with neural networks is to use a separate network for generating the targets y_j in the Q-learning update. More precisely, every C updates we clone the network Q to obtain a target network \hat{Q} and use \hat{Q} for generating the Q-learning targets y_j for the following C updates to Q . This modification makes the algorithm more stable compared to standard online Q-learning, where an update that increases $Q(s_t, a_t)$ often also increases $Q(s_{t+1}, a)$ for all a and hence also increases the target y_t , possibly leading to oscillations or divergence of the policy. Generating the targets using an older set of parameters adds a delay between the time an update to Q is made and the time the update affects the targets y_j , making divergence or oscillations much more unlikely.

We also found it helpful to clip the error term from the update $r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s,a; \theta_i)$ to be between -1 and 1 . Because the absolute value loss function $|x|$ has a derivative of -1 for all negative values of x and a derivative of 1 for all positive values of x , clipping the squared error to be between -1 and 1 corresponds to using an absolute value loss function for errors outside of the $(-1, 1)$ interval. This form of error clipping further improved the stability of the algorithm.

Algorithm 1: deep Q-learning with experience replay.

Initialize replay memory D to capacity N

Initialize action-value function Q with random weights θ

Initialize target action-value function \hat{Q} with weights $\theta^- = \theta$

For episode = 1, M **do**

Initialize sequence $s_1 = \{x_1\}$ and preprocessed sequence $\phi_1 = \phi(s_1)$

For $t = 1, T$ **do**

With probability ε select a random action a_t

otherwise select $a_t = \arg\max_a Q(\phi(s_t), a; \theta)$

Execute action a_t in emulator and observe reward r_t and image x_{t+1}

Set $s_{t+1} = s_t, a_t, x_{t+1}$ and preprocess $\phi_{t+1} = \phi(s_{t+1})$

Store transition $(\phi_t, a_t, r_t, \phi_{t+1})$ in D

Sample random minibatch of transitions $(\phi_j, a_j, r_j, \phi_{j+1})$ from D

Set $y_j = \begin{cases} r_j & \text{if episode terminates at step } j+1 \\ r_j + \gamma \max_{a'} \hat{Q}(\phi_{j+1}, a'; \theta^-) & \text{otherwise} \end{cases}$

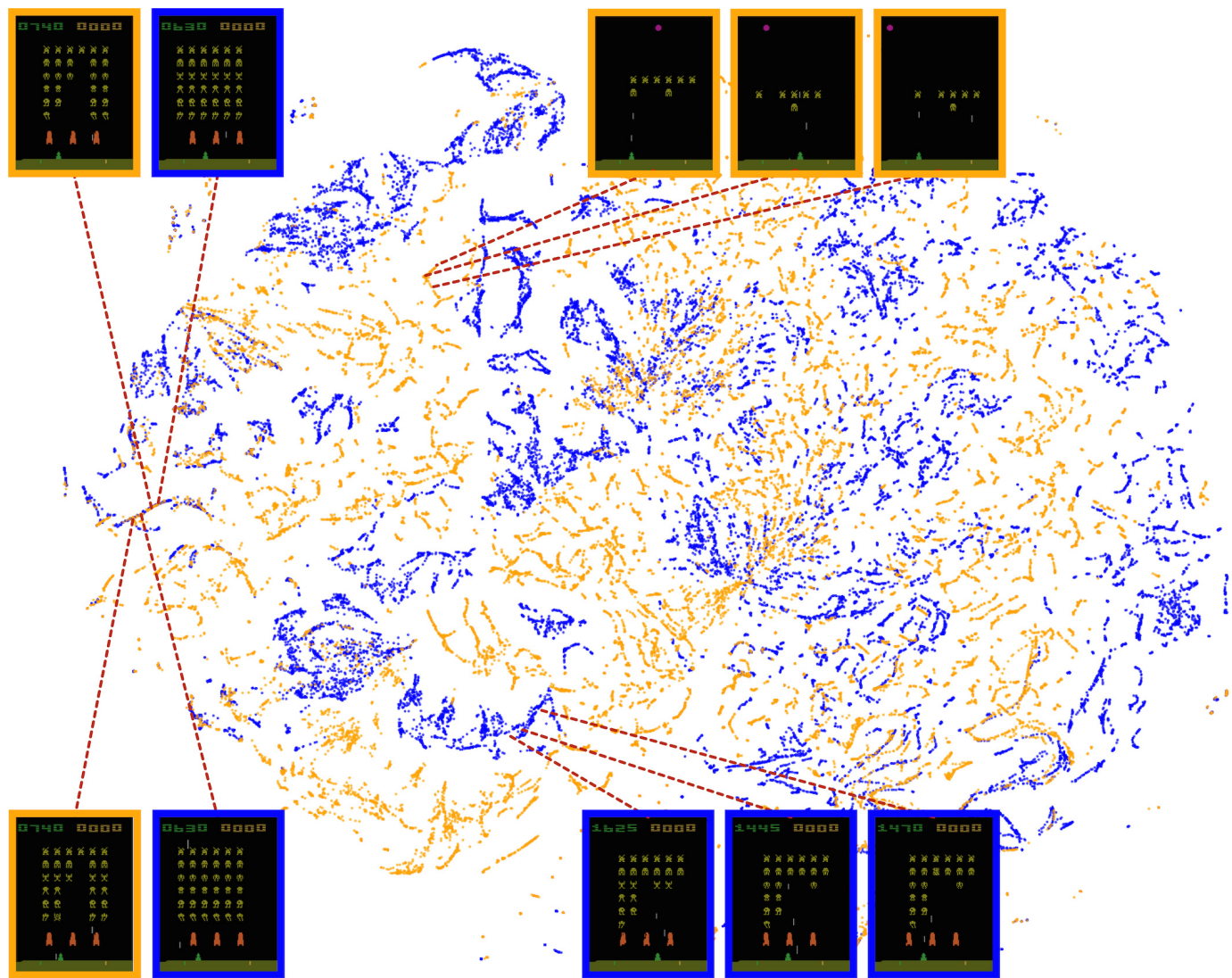
Perform a gradient descent step on $(y_j - Q(\phi_j, a_j; \theta))^2$ with respect to the network parameters θ

Every C steps reset $\hat{Q} = Q$

End For

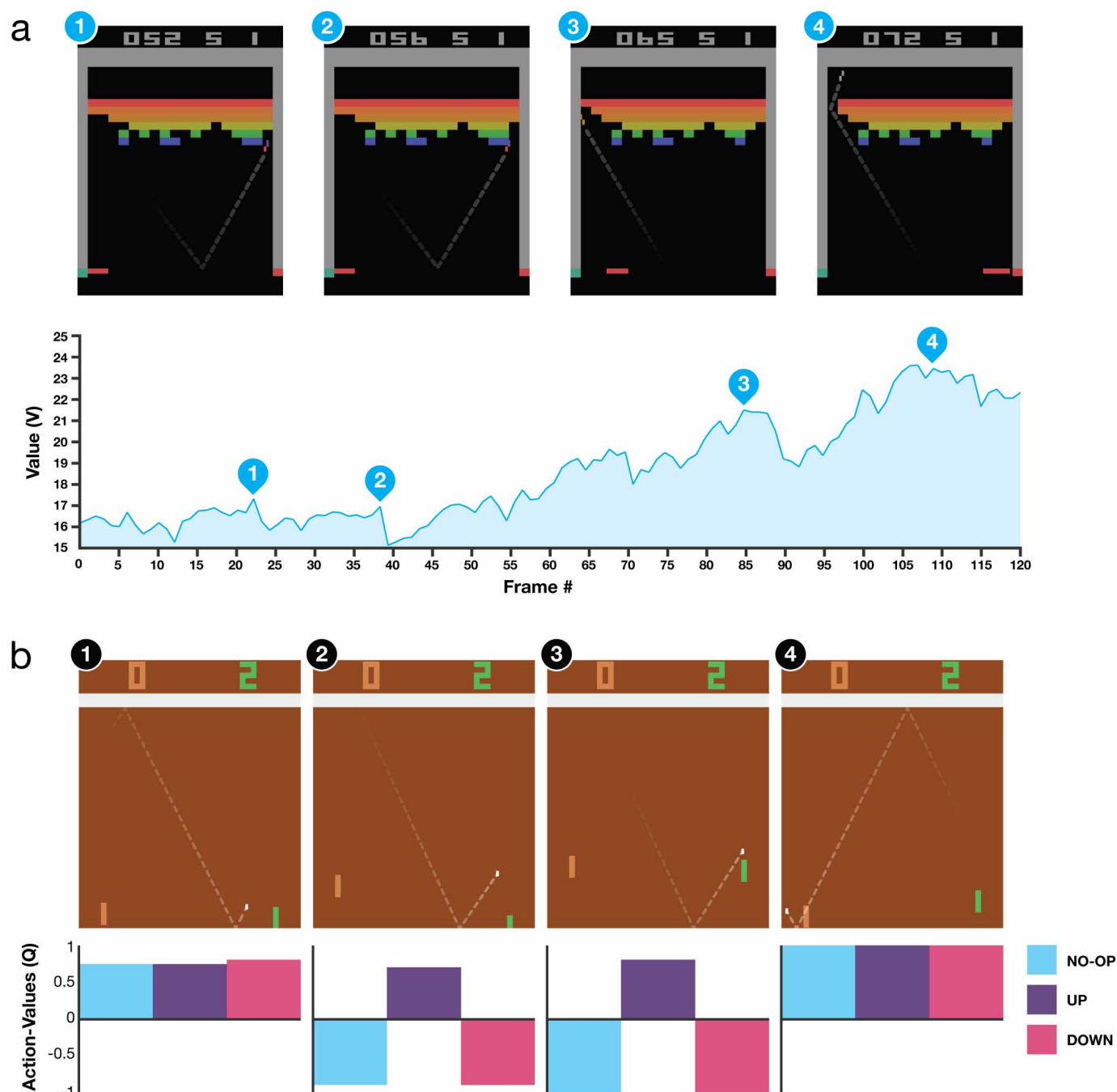
End For

31. Jarrett, K., Kavukcuoglu, K., Ranzato, M. A. & LeCun, Y. What is the best multi-stage architecture for object recognition? *Proc. IEEE. Int. Conf. Comput. Vis.* 2146–2153 (2009).
32. Nair, V. & Hinton, G. E. Rectified linear units improve restricted Boltzmann machines. *Proc. Int. Conf. Mach. Learn.* 807–814 (2010).
33. Kaelbling, L. P., Littman, M. L. & Cassandra, A. R. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* **101**, 99–134 (1994).



Extended Data Figure 1 | Two-dimensional t-SNE embedding of the representations in the last hidden layer assigned by DQN to game states experienced during a combination of human and agent play in Space Invaders. The plot was generated by running the t-SNE algorithm²⁵ on the last hidden layer representation assigned by DQN to game states experienced during a combination of human (30 min) and agent (2 h) play. The fact that there is similar structure in the two-dimensional embeddings corresponding to the DQN representation of states experienced during human play (orange

points) and DQN play (blue points) suggests that the representations learned by DQN do indeed generalize to data generated from policies other than its own. The presence in the t-SNE embedding of overlapping clusters of points corresponding to the network representation of states experienced during human and agent play shows that the DQN agent also follows sequences of states similar to those found in human play. Screenshots corresponding to selected states are shown (human: orange border; DQN: blue border).



Extended Data Figure 2 | Visualization of learned value functions on two games, Breakout and Pong. **a**, A visualization of the learned value function on the game Breakout. At time points 1 and 2, the state value is predicted to be ~ 17 and the agent is clearing the bricks at the lowest level. Each of the peaks in the value function curve corresponds to a reward obtained by clearing a brick. At time point 3, the agent is about to break through to the top level of bricks and the value increases to ~ 21 in anticipation of breaking out and clearing a large set of bricks. At point 4, the value is above 23 and the agent has broken through. After this point, the ball will bounce at the upper part of the bricks clearing many of them by itself. **b**, A visualization of the learned action-value function on the game Pong. At time point 1, the ball is moving towards the paddle controlled by the agent on the right side of the screen and the values of

all actions are around 0.7, reflecting the expected value of this state based on previous experience. At time point 2, the agent starts moving the paddle towards the ball and the value of the 'up' action stays high while the value of the 'down' action falls to -0.9 . This reflects the fact that pressing 'down' would lead to the agent losing the ball and incurring a reward of -1 . At time point 3, the agent hits the ball by pressing 'up' and the expected reward keeps increasing until time point 4, when the ball reaches the left edge of the screen and the value of all actions reflects that the agent is about to receive a reward of 1. Note, the dashed line shows the past trajectory of the ball purely for illustrative purposes (that is, not shown during the game). With permission from Atari Interactive, Inc.

Extended Data Table 1 | List of hyperparameters and their values

Hyperparameter	Value	Description
minibatch size	32	Number of training cases over which each stochastic gradient descent (SGD) update is computed.
replay memory size	1000000	SGD updates are sampled from this number of most recent frames.
agent history length	4	The number of most recent frames experienced by the agent that are given as input to the Q network.
target network update frequency	10000	The frequency (measured in the number of parameter updates) with which the target network is updated (this corresponds to the parameter C from Algorithm 1).
discount factor	0.99	Discount factor gamma used in the Q-learning update.
action repeat	4	Repeat each action selected by the agent this many times. Using a value of 4 results in the agent seeing only every 4th input frame.
update frequency	4	The number of actions selected by the agent between successive SGD updates. Using a value of 4 results in the agent selecting 4 actions between each pair of successive updates.
learning rate	0.00025	The learning rate used by RMSProp.
gradient momentum	0.95	Gradient momentum used by RMSProp.
squared gradient momentum	0.95	Squared gradient (denominator) momentum used by RMSProp.
min squared gradient	0.01	Constant added to the squared gradient in the denominator of the RMSProp update.
initial exploration	1	Initial value of ϵ in ϵ -greedy exploration.
final exploration	0.1	Final value of ϵ in ϵ -greedy exploration.
final exploration frame	1000000	The number of frames over which the initial value of ϵ is linearly annealed to its final value.
replay start size	50000	A uniform random policy is run for this number of frames before learning starts and the resulting experience is used to populate the replay memory.
no-op max	30	Maximum number of "do nothing" actions to be performed by the agent at the start of an episode.

The values of all the hyperparameters were selected by performing an informal search on the games Pong, Breakout, Seaquest, Space Invaders and Beam Rider. We did not perform a systematic grid search owing to the high computational cost, although it is conceivable that even better results could be obtained by systematically tuning the hyperparameter values.

Extended Data Table 2 | Comparison of games scores obtained by DQN agents with methods from the literature^{12,15} and a professional human games tester

Game	Random Play	Best Linear Learner	Contingency (SARSA)	Human	DQN (\pm std)	Normalized DQN (% Human)
Alien	227.8	939.2	103.2	6875	3069 (\pm 1093)	42.7%
Amidar	5.8	103.4	183.6	1676	739.5 (\pm 3024)	43.9%
Assault	222.4	628	537	1496	3359(\pm 775)	246.2%
Asterix	210	987.3	1332	8503	6012 (\pm 1744)	70.0%
Asteroids	719.1	907.3	89	13157	1629 (\pm 542)	7.3%
Atlantis	12850	62687	852.9	29028	85641(\pm 17600)	449.9%
Bank Heist	14.2	190.8	67.4	734.4	429.7 (\pm 650)	57.7%
Battle Zone	2360	15820	16.2	37800	26300 (\pm 7725)	67.6%
Beam Rider	363.9	929.4	1743	5775	6846 (\pm 1619)	119.8%
Bowling	23.1	43.9	36.4	154.8	42.4 (\pm 88)	14.7%
Boxing	0.1	44	9.8	4.3	71.8 (\pm 8.4)	1707.9%
Breakout	1.7	5.2	6.1	31.8	401.2 (\pm 26.9)	1327.2%
Centipede	2091	8803	4647	11963	8309(\pm 5237)	63.0%
Chopper Command	811	1582	16.9	9882	6687 (\pm 2916)	64.8%
Crazy Climber	10781	23411	149.8	35411	114103 (\pm 22797)	419.5%
Demon Attack	152.1	520.5	0	3401	9711 (\pm 2406)	294.2%
Double Dunk	-18.6	-13.1	-16	-15.5	-18.1 (\pm 2.6)	17.1%
Enduro	0	129.1	159.4	309.6	301.8 (\pm 24.6)	97.5%
Fishing Derby	-91.7	-89.5	-85.1	5.5	-0.8 (\pm 19.0)	93.5%
Freeway	0	19.1	19.7	29.6	30.3 (\pm 0.7)	102.4%
Frostbite	65.2	216.9	180.9	4335	328.3 (\pm 250.5)	6.2%
Gopher	257.6	1288	2368	2321	8520 (\pm 3279)	400.4%
Gravitar	173	387.7	429	2672	306.7 (\pm 223.9)	5.3%
H.E.R.O.	1027	6459	7295	25763	19950 (\pm 158)	76.5%
Ice Hockey	-11.2	-9.5	-3.2	0.9	-1.6 (\pm 2.5)	79.3%
James Bond	29	202.8	354.1	406.7	576.7 (\pm 175.5)	145.0%
Kangaroo	52	1622	8.8	3035	6740 (\pm 2959)	224.2%
Krull	1598	3372	3341	2395	3805 (\pm 1033)	277.0%
Kung-Fu Master	258.5	19544	29151	22736	23270 (\pm 5955)	102.4%
Montezuma's Revenge	0	10.7	259	4367	0 (\pm 0)	0.0%
Ms. Pacman	307.3	1692	1227	15693	2311(\pm 525)	13.0%
Name This Game	2292	2500	2247	4076	7257 (\pm 547)	278.3%
Pong	-20.7	-19	-17.4	9.3	18.9 (\pm 1.3)	132.0%
Private Eye	24.9	684.3	86	69571	1788 (\pm 5473)	2.5%
Q*Bert	163.9	613.5	960.3	13455	10596 (\pm 3294)	78.5%
River Raid	1339	1904	2650	13513	8316 (\pm 1049)	57.3%
Road Runner	11.5	67.7	89.1	7845	18257 (\pm 4268)	232.9%
Robotank	2.2	28.7	12.4	11.9	51.6 (\pm 4.7)	509.0%
Seaquest	68.4	664.8	675.5	20182	5286(\pm 1310)	25.9%
Space Invaders	148	250.1	267.9	1652	1976 (\pm 893)	121.5%
Star Gunner	664	1070	9.4	10250	57997 (\pm 3152)	598.1%
Tennis	-23.8	-0.1	0	-8.9	-2.5 (\pm 1.9)	143.2%
Time Pilot	3568	3741	24.9	5925	5947 (\pm 1600)	100.9%
Tutankham	11.4	114.3	98.2	167.6	186.7 (\pm 41.9)	112.2%
Up and Down	533.4	3533	2449	9082	8456 (\pm 3162)	92.7%
Venture	0	66	0.6	1188	380.0 (\pm 238.6)	32.0%
Video Pinball	16257	16871	19761	17298	42684 (\pm 16287)	2539.4%
Wizard of Wor	563.5	1981	36.9	4757	3393 (\pm 2019)	67.5%
Zaxxon	32.5	3365	21.4	9173	4977 (\pm 1235)	54.1%

Best Linear Learner is the best result obtained by a linear function approximator on different types of hand designed features¹². Contingency (SARSA) agent figures are the results obtained in ref. 15. Note the figures in the last column indicate the performance of DQN relative to the human games tester, expressed as a percentage, that is, $100 \times (\text{DQN score} - \text{random play score}) / (\text{human score} - \text{random play score})$.

Extended Data Table 3 | The effects of replay and separating the target Q-network

Game	With replay, with target Q	With replay, without target Q	Without replay, with target Q	Without replay, without target Q
Breakout	316.8	240.7	10.2	3.2
Enduro	1006.3	831.4	141.9	29.1
River Raid	7446.6	4102.8	2867.7	1453.0
Seaquest	2894.4	822.6	1003.0	275.8
Space Invaders	1088.9	826.3	373.2	302.0

DQN agents were trained for 10 million frames using standard hyperparameters for all possible combinations of turning replay on or off, using or not using a separate target Q-network, and three different learning rates. Each agent was evaluated every 250,000 training frames for 135,000 validation frames and the highest average episode score is reported. Note that these evaluation episodes were not truncated at 5 min leading to higher scores on Enduro than the ones reported in Extended Data Table 2. Note also that the number of training frames was shorter (10 million frames) as compared to the main results presented in Extended Data Table 2 (50 million frames).

Extended Data Table 4 | Comparison of DQN performance with linear function approximator

Game	DQN	Linear
Breakout	316.8	3.00
Enduro	1006.3	62.0
River Raid	7446.6	2346.9
Seaquest	2894.4	656.9
Space Invaders	1088.9	301.3

The performance of the DQN agent is compared with the performance of a linear function approximator on the 5 validation games (that is, where a single linear layer was used instead of the convolutional network, in combination with replay and separate target network). Agents were trained for 10 million frames using standard hyperparameters, and three different learning rates. Each agent was evaluated every 250,000 training frames for 135,000 validation frames and the highest average episode score is reported. Note that these evaluation episodes were not truncated at 5 min leading to higher scores on Enduro than the ones reported in Extended Data Table 2. Note also that the number of training frames was shorter (10 million frames) as compared to the main results presented in Extended Data Table 2 (50 million frames).

Evolution of the new vertebrate head by co-option of an ancient chordate skeletal tissue

David Jandzik^{1,2*}, Aaron T. Garnett^{1*}, Tyler A. Square¹, Maria V. Cattell¹, Jr-Kai Yu³ & Daniel M. Medeiros¹

A defining feature of vertebrates (craniates) is a pronounced head that is supported and protected by a robust cellular endoskeleton. In the first vertebrates, this skeleton probably consisted of collagenous cellular cartilage, which forms the embryonic skeleton of all vertebrates and the adult skeleton of modern jawless and cartilaginous fish. In the head, most cellular cartilage is derived from a migratory cell population called the neural crest, which arises from the edges of the central nervous system. Because collagenous cellular cartilage and neural crest cells have not been described in invertebrates¹, the appearance of cellular cartilage derived from neural crest cells is considered a turning point in vertebrate evolution². Here we show that a tissue with many of the defining features of vertebrate cellular cartilage transiently forms in the larvae of the invertebrate chordate *Branchiostoma floridae* (Florida amphioxus). We also present evidence that during evolution, a key regulator of vertebrate cartilage development, *SoxE*, gained new *cis*-regulatory sequences that subsequently directed its novel expression in neural crest cells. Together, these results suggest that the origin of the vertebrate head skeleton did not depend on the evolution of a new skeletal tissue, as is commonly thought, but on the spread of this tissue throughout the head. We further propose that the evolution of *cis*-regulatory elements near an ancient regulator of cartilage differentiation was a major factor in the evolution of the vertebrate head skeleton.

The histological properties of cellular cartilage in larval and adult vertebrates have been studied for over a century, and many specialized cartilage subtypes have been identified³. In contrast to the diversity of

cartilage in adults, at embryonic stages the head skeletons of all living vertebrates consist of a single type of histologically distinct cellular cartilage. This embryonic cartilage consists of tightly packed polygonal or disc-shaped cells that secrete a thin, homogeneous extracellular matrix material composed of fibrillar collagen and chondroitin sulphate proteoglycans³. While classical and modern histological examinations have not identified a clear homologue of vertebrate cellular cartilage in invertebrates, a few invertebrates have cellular-cartilage-like endoskeletal elements. For example, among the protostomes, horseshoe crabs (Merostomata), cephalopod molluscs (Cephalopoda) and sabellid polychaete worms (Sabellidae) have cartilage-like tissues⁴, although their phylogenetic distribution suggests that these tissues evolved independently of vertebrate cartilage, and of each other. Among the deuterostomes, both hemichordates and cephalochordates (amphioxus) have stiff, acellular pharyngeal endoskeletons that incorporate fibrillar collagen⁵. In addition, amphioxus has an oral skeleton that supports its tentacles (cirri) and that forms during metamorphosis (Fig. 1a–e). Although this oral skeleton is ensheathed in a thick integument that does not contain fibrillar collagen, recent work has shown that it has a cellularized core^{6,7}. Unlike other deuterostomes, urochordates lack rigid endoskeletal elements⁴, despite their status as the sister group to vertebrates.

A recurring theme in evolutionary developmental biology is how structures in distantly related taxa can arise via conserved developmental mechanisms, revealing unexpected homology. We decided to test whether, despite its unusual histology in adults, the oral skeleton of amphioxus develops via mechanisms that are homologous to those of vertebrate

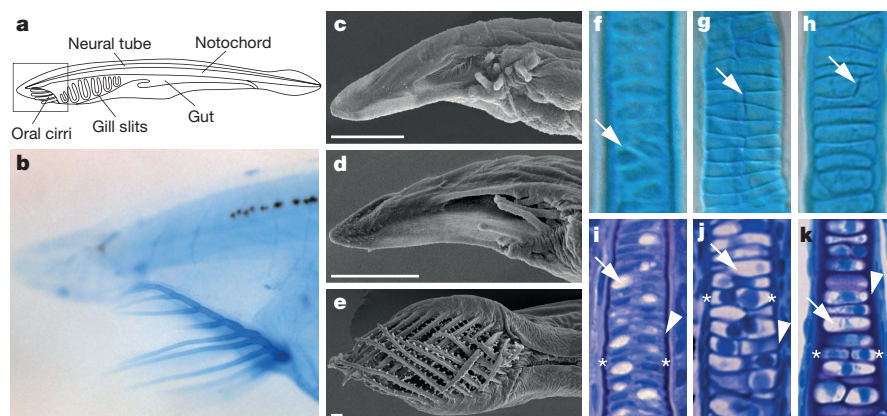


Figure 1 | Development of the amphioxus oral skeleton. **a**, A metamorphic amphioxus larva. The oral region, shown in **b–e**, is boxed. **b**, The oral skeleton of a metamorphosed larva stained with alcian blue. **c**, **d**, Scanning electron micrographs of forming cirri in early metamorphic (**c**) and mid-metamorphic (**d**) larvae. Scale bar, 100 μ m. **e**, A scanning electron micrograph of adult oral cirri. Scale bar, 100 μ m. **f–h**, Alcian blue staining of chondrocytes in a larval amphioxus cirrus (**f**), a larval *Petromyzon marinus* gill bar (**g**) and a larval zebrafish gill bar (**h**). The arrows point to chondrocytes that are

dividing perpendicular to the axis of growth and intercalating to form stacks of discoidal cells. **i–k**, Toluidine blue staining of a larval amphioxus cirrus (**i**), a larval lamprey gill bar (**j**) and a larval zebrafish gill bar (**k**). The nuclei are blue, and the acidic extracellular matrix of the cellular cartilage is purple. The asterisks indicate the paired nuclei of dividing chondrocytes. The arrows point to vacuoles in maturing chondrocytes. The arrowheads indicate the acidic extracellular matrix. Original magnification, $\times 1,000$ (**f**, **i**), $\times 700$ (**g**, **j**) and $\times 500$ (**h**, **k**).

¹Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado 80309, USA. ²Department of Zoology, Comenius University, Bratislava 84215, Slovakia. ³Institute of Cellular and Organismic Biology, Academia Sinica, Taipei 11529, Taiwan.

*These authors contributed equally to this work.

cellular cartilage. Because the amphioxus oral skeleton forms at the end of an extended planktonic larval phase that is difficult to obtain in the field, its development has never been described. Using new methods for the continuous laboratory culture of the Florida amphioxus (*B. floridae*), together with a protocol for artificially inducing synchronized metamorphosis⁸, we were able to obtain live amphioxus larvae at various stages of oral skeleton development. We first stained these larvae with the classic histological stain for vertebrate cellular cartilage, alcian blue, and found strong specific reactivity with the nascent oral skeleton (Fig. 1b, f). Closer histological analysis of individual skeletal rods at this stage revealed discoidal cells with large vacuoles that were dividing perpendicular to the axis of cirrus growth and were surrounded by an acidic extracellular matrix (Fig. 1f, i). This histology is highly similar to that of the gill bar cartilage in embryonic lampreys (*Petromyzon marinus*) and zebrafish (*Danio rerio*) (Fig. 1g, h, j, k).

We next asked whether the development of the oral skeleton in amphioxus requires the same intercellular signalling pathways as the cellular cartilage in vertebrates. Fibroblast growth factor (FGF)-mediated signalling is a conserved essential regulator of cellular cartilage differentiation in both jawed and jawless vertebrates^{9,10}. We thus exposed metamorphosing amphioxus larvae to SU5402 and UO126 (refs 11, 12), inhibitors of FGF-mediated signalling that block cellular cartilage differentiation in vertebrates. Treatment with either inhibitor suppressed the formation of the oral skeleton in metamorphosing amphioxus larvae (Fig. 2a–c and Extended Data Fig. 1c–h). Importantly, these inhibitors did not generally inhibit development, as the treated larvae displayed other temporally appropriate signs of metamorphosis, including formation of

the secondary gill bars and metapleural folds and ventral migration of the mouth⁸.

A defining feature of vertebrate cellular cartilage is the secretion of fibrillar collagen, the major protein component of its extracellular matrix. Although previous histological examinations of amphioxus failed to find evidence of collagen fibres in oral cirri⁶, amputation of adult tentacles leads to the transcription of fibrillar collagen and *SoxE* messenger RNA at the site of amputation⁷. We assayed amphioxus fibrillar collagen (*ColA*) expression during metamorphosis, when the oral skeleton is differentiating. We found intense expression of *ColA* in the chondrocytes of nascent oral cirri and in the mesothelium that lines the anterior coeloms (Fig. 2d and Extended Data Fig. 2b, g–i).

In the cellular cartilage of jawed vertebrates, fibrillar collagen expression and chondrocyte differentiation are activated by the transcription factor SOX9, a SOXE family member and key regulator of chondrogenesis¹³. The binding of SOX9 to the promoter of the gene encoding fibrillar collagen, *Col2a1*, is facilitated by transcription factors of the *SoxD* subfamily¹⁴, while the expression of *Sox9* depends on FGF-mediated signalling through FGF receptors (FGFRs)¹⁵. In lampreys, *SoxE* genes¹⁶ and FGF-mediated signalling through FGFRs¹⁰ are also required for the proper differentiation of neural crest cell (NCC)-derived chondrocytes, while *SoxD* is expressed in prechondrocytes¹⁷. These results suggest that SOXE, SOXD and FGF signalling are ancient core components of the vertebrate cartilage gene program. We thus assessed whether the developing amphioxus oral skeleton expresses *SoxE*, *SoxD* and FGF signalling pathway components. We detected *SoxE*, *SoxD* and *FGFR* mRNA in oral chondrocytes and the surrounding mesothelium (Fig. 2e, f and Extended Data Fig. 2c, d). We observed similar expression of *Ets* and *DUSP6/7/9*, homologues of two FGF-mediated signalling target genes expressed in chondrogenic NCCs in zebrafish, mice and *Xenopus laevis*^{18,19} (Fig. 2g and Extended Data Fig. 2e, f, j, k). Taken together, our data show that the developing amphioxus oral skeleton displays the core, conserved histological, developmental and molecular features of vertebrate embryonic cellular cartilage.

Most of the cellular cartilage in the vertebrate head is derived from NCCs. Amphioxus lacks NCCs and probably forms its oral skeleton from the mesothelium that lines the anterior coeloms²⁰. This implies that the genetic program for generating cellular cartilage was primitively deployed in mesendoderm and later recruited by NCCs. The repurposing of ancient genes and genetic programs is a recognized way in which novelty arises during evolution. However, it is unresolved whether such co-option events are typically driven by changes in *cis*-regulatory sequences, changes in the function of transcription factors or a combination of the two. In all vertebrates that have been examined, including lampreys, SOXE transcription in NCCs is activated by the transcription factor TFAP2 and is maintained by auto-regulation between SOXE family members^{21–24}. These regulatory interactions are direct and involve the physical binding of TFAP2 and SOXE paralogues to SOXE enhancers in NCCs^{21,22}. Previous work has shown that, despite new roles in NCC development, neither TFAP2 nor SOXE has acquired new DNA-binding properties in vertebrates^{22,25}. This implies that the novel expression of *SoxE* in NCCs was driven by changes in *SoxE cis*-regulatory sequences rather than by changes in transcription factor function. To directly test this hypothesis, we performed an interspecific assay of *SoxE cis* regulation. To ensure that all of the relevant amphioxus *cis*-regulatory sequences were queried, we built a bacterial artificial chromosome (BAC) reporter construct that contains the entire amphioxus *SoxE* locus and several flanking genes, and we tested this reporter in zebrafish (Fig. 3b). We reasoned that if changes in *SoxE cis* regulation were involved in the recruitment of *SoxE* to NCCs, then the reporter should recapitulate the expression pattern of endogenous *SoxE* in amphioxus (Fig. 3a). Alternatively, if *SoxE cis* regulation has been conserved across chordates, the reporter should be active in NCCs, as shown previously for amphioxus *Hox* enhancers²⁶. We observed reporter activity transiently throughout the neural tube and tail-bud region in early neurulae but not in migrating NCCs or chondrocytes at any developmental stage (Fig. 3c–f and

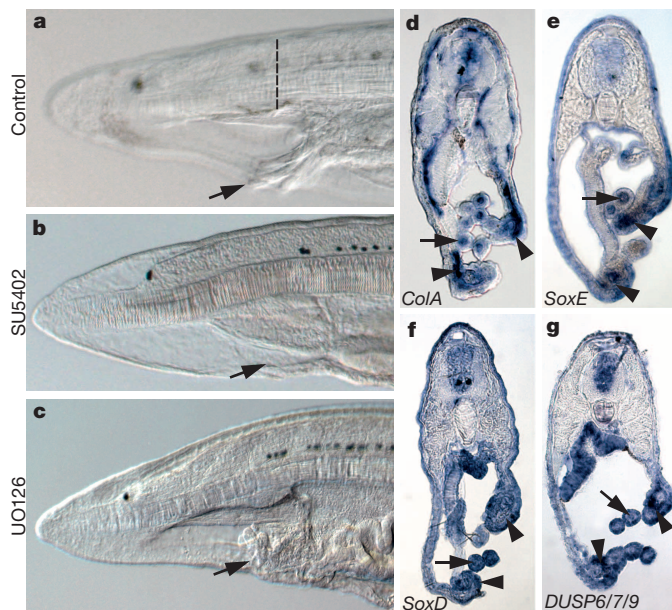


Figure 2 | The amphioxus oral skeleton requires FGF-mediated signalling for formation and expresses orthologues of vertebrate cartilage markers. **a**, A phase contrast image of cirri in a control larva treated with T3 thyroid hormone, which induces metamorphosis in amphioxus. Ninety-six per cent of T3-treated control larvae (22 of 23, from two experiments) developed a normal oral skeleton (arrow) after 4–5 days, as previously reported⁸. **In situ** hybridization (**d–g**) was performed on sections at the level of the dashed line. **b**, **c**, Representative phase contrast images of larvae treated with T3 and 50 μM SU5402 (**b**) or T3 and 10 μM UO126 (**c**). All SU5402-treated larvae (26 of 26, from two experiments) and 83% of UO126-treated larvae (10/12, from one experiment) lacked oral cirri (arrows) but displayed other signs of metamorphosis, including metapleural folds, ventral mouth migration and secondary gill bar formation. **d–g**, Expression of amphioxus fibrillar collagen (*ColA*) (**d**), *SoxE* (**e**), *SoxD* (**f**) and *DUSP6/7/9* (**g**) mRNA (shown in blue), as determined by *in situ* hybridization, in the oral region of metamorphic amphioxus larvae. Expression in chondrocytes is indicated by arrows, and expression in the mesothelium is indicated by arrowheads.

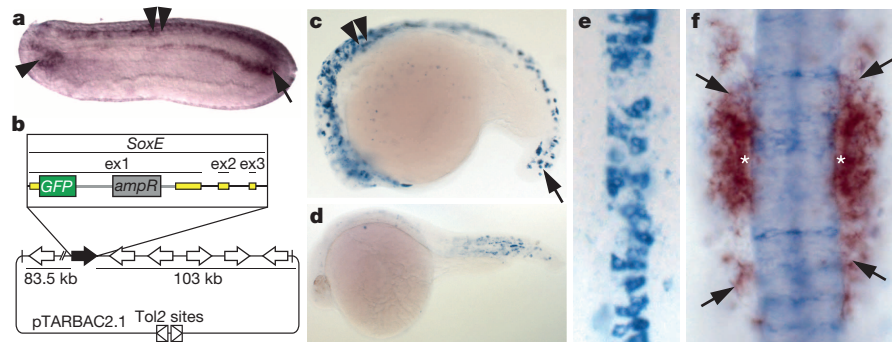


Figure 3 | A reporter construct incorporating the amphioxus *SoxE* locus recapitulates the amphioxus *SoxE* expression pattern in zebrafish embryos.

a, Transient expression of amphioxus *SoxE* transcripts (purple) in scattered neural tube cells (double arrowhead), the posterior axial mesoderm including the tail bud (arrow) and the anterior mesoderm (arrowhead) in a 20-h, 12-somite neurula, as previously described²⁷. **b**, Schematic of the amphioxus *SoxE* reporter construct. The construct spans the *SoxE* locus and 186 kilobases (kb) of flanking sequence, including adjacent genes, and is roughly equivalent to 1.3 million base pairs of human genomic sequence. The gene encoding green fluorescent protein (*GFP*) was fused in frame with the first exon (*ex1*), and Tol2 recombination arms were added to facilitate genomic integration. *ampR*, ampicillin resistance. **c**, A zebrafish embryo injected at the one-cell stage with the amphioxus *SoxE* BAC reporter construct and probed for *GFP* mRNA at the 16-h, 15-somite stage. Mosaic transcription of *GFP* (blue) was consistently observed in the neural tube (double arrowhead) and tail-bud region (arrow). **d**, A zebrafish embryo injected at the one-cell stage with the

amphioxus *SoxE* BAC reporter and probed for *GFP* mRNA at the 22-h, 26+ -somite stage. Like endogenous amphioxus *SoxE*, the *SoxE* BAC reporter expression was transient and was extinguished by the late neurula stages, when vertebrate *SOXE* genes mark migrating and post-migratory NCCs. **e**, A dorsal close-up view of the embryo in **c** at the level of the hindbrain. Activity of the amphioxus *SoxE* BAC reporter occurred throughout the neural tube but not in the surrounding mesenchyme, which included some early migrating NCCs. The image is a composite of three photographs of the same embryo taken at different focal planes. **f**, Double *in situ* hybridization for zebrafish *sox10* (a *SoxE* co-orthologue) (red) and *GFP* (blue) transcripts in an 18-somite zebrafish injected with the *SoxE* BAC reporter construct. The image is a composite of three photographs of the same embryo taken at different focal planes. *GFP* expression was limited to the neural tube and did not overlap with *sox10* expression in the otic placode (asterisks) and migrating NCCs (arrows).

Extended Data Fig. 3). This pattern closely resembles endogenous amphioxus *SoxE* expression in the embryonic neural tube and tail-bud²⁷ (Fig. 3a), regions that do not co-express *Tfap2* and *SoxE*. These results, and the lack of endogenous *soxE* in the lateral and ventral neural tube of zebrafish embryos, suggest that the amphioxus *SoxE* reporter and, by extension, endogenous amphioxus *SoxE* are not regulated by *Tfap2* or *SoxE* proteins.

To verify this finding, we assessed whether the reporter functioned properly in zebrafish embryos with depleted *tfap2* function. While

TFAP2A and *TFAP2C* dual knockdown eliminates the expression of *sox10* and *sox9b* (amphioxus *SoxE* co-orthologues) in zebrafish^{22,28}, the activity of the amphioxus *SoxE* reporter was similar in wildtype zebrafish and in *tfap2a* and *tfap2c* dual knockdown morphants (Extended Data Fig. 3). In the context of previous work, these data suggest that amphioxus and vertebrate *SoxE* *cis* regulation diverged, in part, through the evolution of new non-coding *cis*-regulatory sequences of *tfap2* and *SoxE*. Whether these sequence changes were accompanied by changes in the function of other transcription factors is unclear from our assay.

Our results suggest that the nascent oropharyngeal skeleton of early chordates incorporated collagenous cellular cartilage that is strikingly similar to vertebrate cartilage. In amphioxus, this tissue supports the oral tentacles, which prevent the ingestion of large particles during filter feeding and burrowing. It is likely that similar structures were present in the invertebrate ancestor of the vertebrates, as oral tentacles are found in lamprey larvae, adult hagfish and the chordate fossil *Haikouella lanceolata*²⁹. We posit that the evolution of the vertebrate head skeleton involved two major developmental changes: the spread of collagenous cellular cartilage from the oral region into the pharynx and the head, and the novel differentiation of cellular cartilage from NCCs (Fig. 4). Our data also suggest that the acquisition of chondrogenic potential by NCCs involved the evolution of new transcription-factor-binding sites

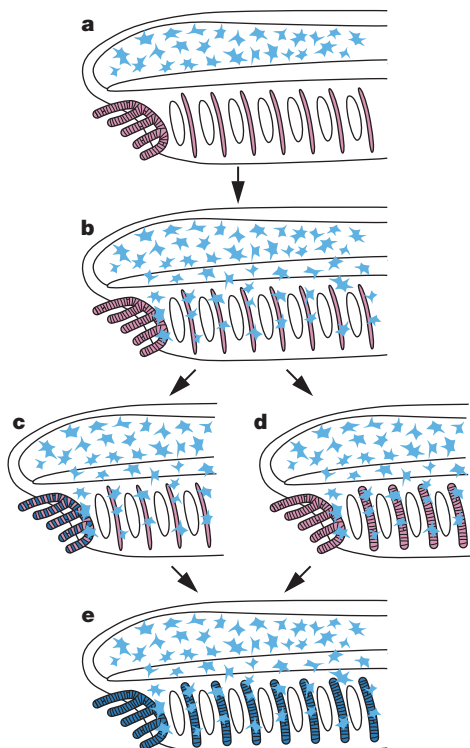


Figure 4 | The evolution of the vertebrate head skeleton via co-option of an ancient cellular cartilage gene program.

a, A hypothetical early chordate with an oral skeleton consisting of mesoderm-derived cellular cartilage (pink polygonal cells) and a pharyngeal skeleton of acellular cartilage (pink rods). *SoxE* controls cellular cartilage differentiation in the oral region and has an unrelated function in a subset of central nervous system precursors (blue stellate cells). **b**, An early pre-vertebrate chordate with migratory non-skeletogenic proto-NCCs expressing *SoxE* (blue stellate cells in the oropharyngeal region). **c**, Exposure to intercellular signals in the oral region activates the cellular cartilage gene program in *SoxE*-expressing proto-NCCs (blue polygonal cells). **d**, Alternatively, mesoderm-derived cellular cartilage (pink polygonal cells) spreads throughout the pharynx, before being replaced by NCC-derived cellular cartilage (**e**). **e**, Subsequent to **c** or **d**, invasion of the pharynx by chondrogenic NCC-derived cellular cartilage gives rise to the NCC-derived head skeleton of vertebrates (blue polygonal cells), similar to the model proposed by Rychel and Swalla⁵.

at the *SoxE* locus. In addition to the oral cartilage, *SoxE* is expressed in the embryonic central nervous system of amphioxus. One of the first steps in the evolution of NCC-derived cellular cartilage may have been *cis*-regulatory mutations that maintained *SoxE* expression in non-skeletogenic proto-NCCs emerging from the central nervous system. Later, when these *SoxE*-expressing migratory cells mixed with mesoderm in the oral region, exposure to skeletogenic signals such as FGFs could have induced their differentiation into chondrocytes (Fig. 4). Thus, the dual roles of the ancestral chordate *SoxE* gene in neural and skeletal development may have predisposed evolving NCCs towards acquiring chondrogenic ability, potentiating the evolution of the vertebrate “new head”²².

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 September; accepted 24 October 2014.

Published online 8 December 2014.

- Medeiros, D. M. The evolution of the neural crest: new perspectives from lamprey and invertebrate neural crest-like cells. *Wiley Interdiscip. Rev. Dev. Biol.* **2**, 1–15 (2013).
- Gans, C. & Northcutt, R. G. Neural crest and the origin of vertebrates: a new head. *Science* **220**, 268–273 (1983).
- Witten, P. E., Huysseune, A. & Hall, B. K. A practical approach for the identification of the many cartilaginous tissues in teleost fish. *J. Appl. Ichthyology* **26**, 257–262 (2010).
- Cole, A. G. & Hall, B. K. The nature and significance of invertebrate cartilages revisited: distribution and histology of cartilage and cartilage-like tissues within the Metazoa. *Zoology (Jena)* **107**, 261–273 (2004).
- Rychel, A. L. & Swalla, B. J. Development and evolution of chordate cartilage. *J. Exp. Zool.* **308B**, 325–335 (2007).
- Wright, G. M., Keeley, F. W. & Robson, P. The unusual cartilaginous tissues of jawless craniates, cephalochordates and invertebrates. *Cell Tissue Res.* **304**, 165–174 (2001).
- Kaneto, S. & Wada, H. Regeneration of amphioxus oral cirri and its skeletal rods: implications for the origin of the vertebrate skeleton. *J. Exp. Zool.* **316B**, 409–417 (2011).
- Paris, M. *et al.* Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr. Biol.* **18**, 825–830 (2008).
- Walshe, J. & Mason, I. Fgf signalling is required for formation of cartilage in the head. *Dev. Biol.* **264**, 522–536 (2003).
- Jandzik, D. *et al.* Roles for FGF in lamprey pharyngeal pouch formation and skeletogenesis highlight ancestral functions in the vertebrate head. *Development* **141**, 629–638 (2014).
- Mohammadi, M. *et al.* Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science* **276**, 955–960 (1997).
- Favata, M. F. *et al.* Identification of a novel inhibitor of mitogen-activated protein kinase kinase. *J. Biol. Chem.* **273**, 18623–18632 (1998).
- Bi, W. M., Deng, J. M., Zhang, Z. P., Behringer, R. R. & de Crombrughe, B. Sox9 is required for cartilage formation. *Nature Genet.* **22**, 85–89 (1999).
- Lefebvre, V., Li, P. & de Crombrughe, B. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *EMBO J.* **17**, 5718–5733 (1998).
- Kumar, M., Ray, P. & Chapman, S. C. Fibroblast growth factor and bone morphogenetic protein signaling are required for specifying prechondrogenic identity in neural crest-derived mesenchyme and initiating the chondrogenic program. *Dev. Dyn.* **241**, 1091–1103 (2012).
- McCauley, D. W. & Bronner-Fraser, M. Importance of SoxE in neural crest development and the evolution of the pharynx. *Nature* **441**, 750–752 (2006).
- Ohtani, K. *et al.* Expression of Sox and fibrillar collagen genes in lamprey larval chondrogenesis with implications for the evolution of vertebrate cartilage. *J. Exp. Zool.* **310B**, 596–607 (2008).
- Znosko, W. A. *et al.* Overlapping functions of Pea3 ETS transcription factors in FGF signaling during zebrafish development. *Dev. Biol.* **342**, 11–25 (2010).
- Li, C. Y., Scott, D. A., Hatch, E., Tian, X. Y. & Mansour, S. L. Dusp6 (Mkp3) is a negative feedback regulator of FGF-stimulated ERK signaling during mouse development. *Development* **134**, 167–176 (2007).
- Conklin, E. G. The embryology of amphioxus. *J. Morphol.* **54**, 69–151 (1932).
- Betancur, P., Bronner-Fraser, M. & Sauka-Spengler, T. Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *Proc. Natl Acad. Sci. USA* **107**, 3570–3575 (2010).
- Van Otterloo, E. *et al.* Novel Tfp2-mediated control of soxE expression facilitated the evolutionary emergence of the neural crest. *Development* **139**, 720–730 (2012).
- Lakiza, O., Miller, S., Bunce, A., Lee, E. M. J. & McCauley, D. W. SoxE gene duplication and development of the lamprey branchial skeleton: insights into development and evolution of the neural crest. *Dev. Biol.* **359**, 149–161 (2011).
- Sauka-Spengler, T., Meulemans, D., Jones, M. & Bronner-Fraser, M. Ancient evolutionary origin of the neural crest gene regulatory network. *Dev. Cell* **13**, 405–420 (2007).
- Cossais, F. *et al.* Replacement of mouse Sox10 by the *Drosophila* ortholog Sox100B provides evidence for co-option of SoxE proteins into vertebrate-specific gene-regulatory networks through altered expression. *Dev. Biol.* **341**, 267–281 (2010).
- Manzanares, M. *et al.* Conservation and elaboration of Hox gene regulation during evolution of the vertebrate head. *Nature* **408**, 854–857 (2000).
- Meulemans, D. & Bronner-Fraser, M. Insights from amphioxus into the evolution of vertebrate cartilage. *PLoS ONE* **2**, e787 (2007).
- Li, W. & Cornell, R. A. Redundant activities of Tfp2a and Tfp2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Dev. Biol.* **304**, 338–354 (2007).
- Chen, J. Y., Huang, D. Y. & Li, C. W. An early Cambrian craniate-like chordate. *Nature* **402**, 518–522 (1999).

Acknowledgements We thank D. Brunelle, C.-H. Tung and T.-K. Huang for assistance with amphioxus husbandry, and P. Tsai and D. W. Stock for use of their microtomes. D.M.M., D.J., T.A.S. and M.V.C. were supported by National Science Foundation grants IOS 1257040 and IOS 1160733 (D.M.M.) and University of Colorado, Boulder start-up funds (D.M.M.). A.T.G. was supported by National Science Foundation grant DBI 0905991. J.-K.Y. was supported by National Science Council Taiwan grants NSC101-2923-B-001-004-MY2 and NSC102-2311-B-001-011-MY3 and Career Development Award AS-98-CDA-L06 from Academia Sinica Taiwan.

Author Contributions D.M.M. designed the project and wrote the manuscript. D.J., A.T.G. and T.A.S. performed the experiments and prepared the images; D.J. assembled the figures; and M.V.C. and J.-K.Y. provided materials and reagents. All authors discussed the results and provided input on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.M.M. (Daniel.Medeiros@Colorado.edu).

METHODS

Amphioxus adults were fed a diet of live unicellular algae and allowed to spawn spontaneously, en masse, in closed aquaria containing 60–150 animals. The larvae from mass spawning events were collected and raised on live unicellular algae and then induced to metamorphose by exposure to T3 thyroid hormone³⁰. Experiments on vertebrate animals were carried out in accordance with the guidelines of the University of Colorado, Boulder, Institutional Animal Care and Use Committee.

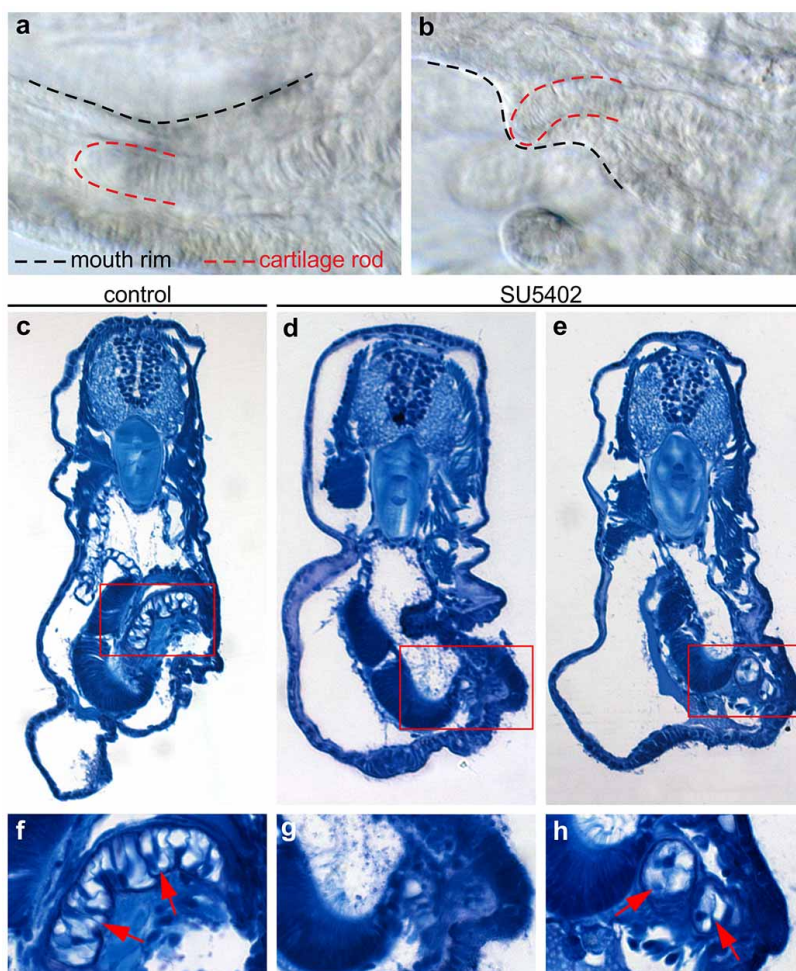
The larvae were fixed, embedded in paraffin, sectioned to 10 µm and *in situ* hybridized to riboprobes using previously reported methods³¹. The *Ets*, *ColA*, *SoxE* and *SoxD* riboprobes were made as previously described²⁷. The *DUSP6/7/9* and *DUSP1/4/5* riboprobes were synthesized using cDNA clones GC035f15 and GC034j09, respectively, from the *B. floridae* Gene Collection Release 1 (ref. 32). The *FGFR* riboprobe was synthesized using cDNA clone CAXG1602 from the *B. floridae* cDNA library, CAXG (NCBI EST accession, FE584207)³³.

Alcian blue and toluidine blue staining of 4-µm sections embedded in JB-4 plastic resin (Polysciences) and scanning electron microscopy were performed according to standard protocols. SU5402 and UO126 (Tocris Bioscience) were dissolved in dimethyl sulphoxide before dilution in seawater for larval treatments. The control and experimental larvae were selected at random from the same mass spawnings. The scoring of treated and control larvae for oral cirrus development was not blinded.

The BAC recombination-mediated genetic engineering was carried out as previously described³⁴, with modifications as follows. A region surrounding the transcription start site of *SoxE* was amplified using the primers 5'-GCATggcgccgGAGGATAGGTACCTATCCGTC-3' and 5'-GCATccatggGCACCAGCGTCCAGTCGTACC-3', digested with *AscI* and *NcoI* and ligated into pLD53.SC2. The resultant plasmid was used to insert *GFP* in frame with *SoxE* in a BAC from the CHORI-302 library (BACPAC Resources Center, accession CH302-98N3; <https://bacpac.chori.org>). Tol2 sites were then inserted into the pTARBAC2.1 backbone. To do this, a plasmid was constructed by inserting the kanamycin resistance gene and a 441-base-pair-targeting sequence from the pTARBAC2.1 backbone into the plasmid pDest Tol2pa using Gateway cloning³⁵. The primers 5'-GCATactagtAGAGGT

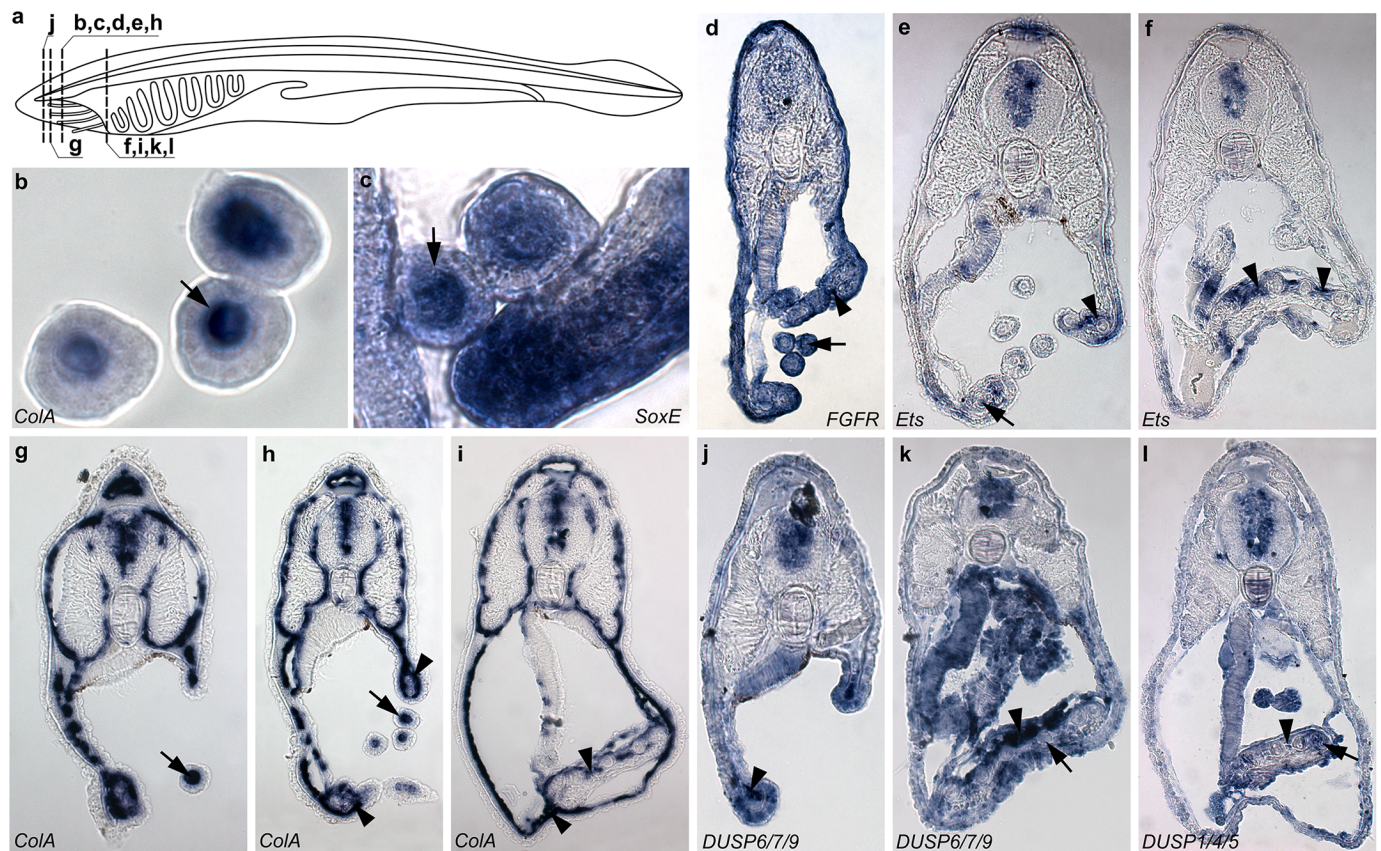
TTGTCCAGGAGTTC-3' and 5'-GCATactagtGCTGGGCTTGCTGAAGGTA GG-3' were used to amplify a cassette from this plasmid containing kanamycin resistance and the pTARBAC2.1-targeting sequence flanked by Tol2 recognition sites. The resultant fragment was digested with *SpeI* and circularized using DNA ligase. Then, bacteria containing the *SoxE* BAC and the pSV1-RecA plasmid were transformed with this circular DNA. Bacteria containing the recombinant BAC were then selected. The resultant reporter was then coinjected with mRNA encoding Tol2 recombinase³⁶ into AB strain zebrafish embryos at the one-to-two cell stage. For some embryos, this injection mix included splice-blocking morpholino antisense oligonucleotides (MOs) against *tfap2a* and *tfap2c* (*tfap2a+c* MO)²⁸. Injected embryos expressing high levels of GFP protein, as revealed by fluorescence microscopy, were then fixed and *in situ* hybridized to *GFP* and *sox10* probes. Several embryos that were coinjected with *tfap2a+c* MO and the BAC reporter were allowed to develop past the 18-somite stage to verify the high penetrance (16/18) reported for the *tfap2a+c* morphant phenotype²⁸.

30. Paris, M. *et al.* Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr. Biol.* **18**, 825–830 (2008).
31. O'Neill, P., McCole, R. B. & Baker, C. V. H. A molecular analysis of neurogenic placode and cranial sensory ganglion development in the shark, *Scyliorhinus canicula*. *Dev. Biol.* **304**, 156–181 (2007).
32. Yu, J. K. *et al.* A cDNA resource for the cephalochordate amphioxus *Branchiostoma floridae*. *Dev. Genes Evol.* **218**, 723–727 (2008).
33. Putnam, N. H. *et al.* The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**, 1064–1071 (2008).
34. Gong, S. C., Kus, L. & Heintz, N. Rapid bacterial artificial chromosome modification for large-scale mouse transgenesis. *Nature Protocols* **5**, 1678–1696 (2010).
35. Villefranc, J. A., Amigo, J. & Lawson, N. D. Gateway compatible vectors for analysis of gene function in the zebrafish. *Dev. Dyn.* **236**, 3077–3087 (2007).
36. Kawakami, K. *et al.* A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev. Cell* **7**, 133–144 (2004).
37. Cameron, T. L., Belluoccio, D., Farlie, P. G., Brachvogel, B. & Bateman, J. F. Global comparative transcriptome analysis of cartilage formation *in vivo*. *BMC Dev. Biol.* **9**, 20 (2009).



Extended Data Figure 1 | Amphioxus oral cartilage differentiation is initiated before cirrus outgrowth and requires FGF-mediated signalling. **a, b**, Phase contrast images of a metamorphic amphioxus larva. **a**, The differentiation of cartilage rods occurs first in the rim of the mouth. **b**, The cartilage rods later grow outwards into nascent cirri. **c**, Toluidine blue staining of a JB-4-embedded control larval section showing the oral cartilage rod (red

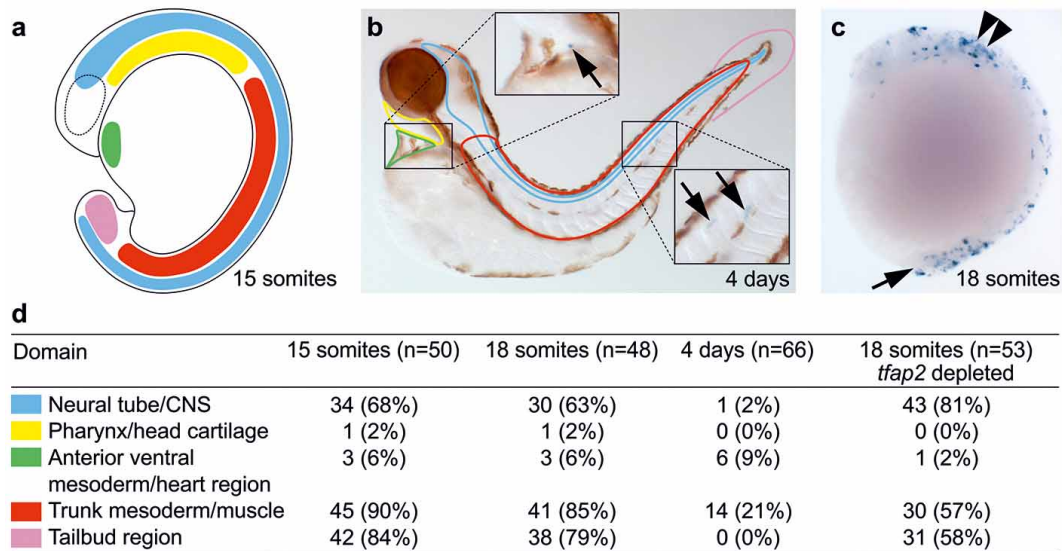
rectangle; arrows in **f**) embedded in the rim of the mouth. **d, e**, Toluidine blue staining of two representative SU5402-treated larvae, sectioned at the same level as the larva in **c**. The differentiation of the oral cartilage bar is completely eliminated in **d** and is strongly reduced in **e**. **f–h**, High magnification views of the sections in **c–e**.



Extended Data Figure 2 | Expression of SoxE, ColA and FGF signalling components in the oral region of metamorphic amphioxus larvae.

a–l, *In situ* hybridization was performed on transverse sections in the planes shown in **a**. **b**, A high magnification view of cirri hybridized with a probe against *ColA*. *ColA* mRNA was detected in the central cartilage rod, which is a single stack of discoidal chondrocytes. **c**, A high magnification view of cirri hybridized with a probe against *SoxE*. *SoxE* mRNA was detected in the central cartilage rod. **d**, A high magnification view of cirri hybridized with a probe

against *FGFR*. **e**, **f**, High magnification views of cirri hybridized with a probe against *Ets*. **g–i**, High magnification views of cirri hybridized with a probe against *ColA*. **j**, **k**, High magnification views of cirri hybridized with a probe against *DUSP6/7/9*. **l**, A high magnification view of cirri hybridized with a probe against *DUSP1/4/5*, an orthologue of mouse *Dusp4*, a gene that is expressed at high levels in mesoderm-derived chondrocytes³⁷. In all panels, the arrows indicate expression in oral chondrocytes, and the arrowheads indicate expression in the associated oral mesothelial cells.



Extended Data Figure 3 | Activity of the amphioxus *SoxE* reporter construct in zebrafish during development. Injected zebrafish embryos displaying broad GFP fluorescence and normal morphology were processed for *in situ* hybridization to detect *GFP* mRNA. Embryos were scored for the expression of GFP in five or more cells in each domain. **a**, A map of the expression domains scored in 15-somite and 18-somite embryos. **b**, A representative 4-day larva showing sporadic expression in the heart region and trunk muscles (arrows in insets). The scored domains are outlined. The image is a composite of eight photographs of the same larva taken at different focal planes. **c**, A representative 18-somite *tfap2a* and *tfap2c* dual knockdown morphant (*tfap2a/c* morphant)²⁸ expressing the amphioxus *SoxE* reporter. The image is a composite of four

photographs of the same embryo taken at different focal planes. The *tfap2a/c* morphants²⁸ displayed a highly penetrant and almost complete loss of NCC marker expression, including *sox10* (a co-orthologue of amphioxus *SoxE*). Like wildtype embryos, *tfap2a/c* morphant embryos displayed mosaic expression of the amphioxus *SoxE* reporter in the neural tube (double arrowheads) and tail bud (arrow). The ability of the amphioxus *SoxE* reporter to function in *tfap2*-depleted embryos supports expression data²⁷ suggesting that amphioxus *SoxE* transcription is *Tfap2* independent. **d**, The numbers and frequencies of embryos and larvae with GFP expression in the indicated domains. The numbers in each column are the pooled results of two to four separate experiments.

Experimentally induced innovations lead to persistent culture via conformity in wild birds

Lucy M. Aplin^{1,2}, Damien R. Farine^{1,3,4}, Julie Morand-Ferron⁵, Andrew Cockburn², Alex Thornton⁶ & Ben C. Sheldon^{1,7}

In human societies, cultural norms arise when behaviours are transmitted through social networks via high-fidelity social learning¹. However, a paucity of experimental studies has meant that there is no comparable understanding of the process by which socially transmitted behaviours might spread and persist in animal populations^{2,3}. Here we show experimental evidence of the establishment of foraging traditions in a wild bird population. We introduced alternative novel foraging techniques into replicated wild sub-populations of great tits (*Parus major*) and used automated tracking to map the diffusion, establishment and long-term persistence of the seeded innovations. Furthermore, we used social network analysis to examine the social factors that influenced diffusion dynamics. From only two trained birds in each sub-population, the information spread rapidly through social network ties, to reach an average of 75% of individuals, with a total of 414 knowledgeable individuals performing 57,909 solutions over all replicates. The sub-populations were heavily biased towards using the technique that was originally introduced, resulting in established local traditions that were stable over two generations, despite a high population turnover. Finally, we demonstrate a strong effect of social conformity, with individuals disproportionately adopting the most frequent local variant when first acquiring an innovation, and continuing to favour social information over personal information. Cultural conformity is thought to be a key factor in the evolution of complex culture in humans^{4–7}. In providing the first experimental demonstration of conformity in a wild non-primate, and of cultural norms in foraging techniques in any wild animal, our results suggest a much broader taxonomic occurrence of such an apparently complex cultural behaviour.

Social learning, in which animals learn from others, can enable novel behaviours to spread between individuals, creating group-level behaviours, including traditions and culture^{6,8,9}. Social transmission occurs between interacting individuals; hence, group dynamics and population structure will determine the spread and persistence of traditions^{2,3,9–11}. Additionally, individuals may strategically use social learning to maximize its adaptive value, with consequences for when, how and what traditions are established^{4,12}. However, while the capacity for social learning has been described in many phylogenetically diverse taxa¹³ and has been detailed in comprehensive laboratory studies^{13–15}, we have little knowledge of the social dynamics associated with such learning in natural systems. Experimentally quantifying cultural transmission in wild populations remains difficult, with limitations associated with isolating and training individuals⁵, tracking the spread of information across large numbers of animals¹⁴ and eliminating alternative explanations such as individual trial-and-error learning^{8,14}.

Early observational studies of tits provide one of the most widely cited examples of animal innovation and culture, when British birds famously began to pierce the foil caps of milk bottles to take the cream^{16–18}. More generally, great tits (*P. major*) are known to be highly innovative, opportunistic foragers¹⁹ and to use social information in a wide range of contexts²⁰.

This life history, coupled with their fission–fusion social structure²¹, makes them excellent models for a large-scale empirical investigation of the social processes associated with cultural transmission. Here we used a novel system that incorporates automated data collection and passive integrated transponder tags, together with recently developed methods for social network analysis, to investigate the spread, establishment and persistence of experimentally seeded traditions in wild great tits.

We first developed an automated puzzle box that is baited with live mealworms (Fig. 1a), and performed a cultural diffusion experiment based on the two-action and control design¹⁴ but where treatment groups were exposed to a demonstrator trained on one of two distinct but equivalent actions. Two resident males were caught from each of eight sub-populations and exposed to one of three training regimens in captivity. In the first condition ('control'), for which there were three sub-populations (that is, three replicates), neither individual was given any training. In the second condition ('option A'; two replicates), both individuals were trained to access food from the puzzle box by using their bill to move the blue side of the sliding door from left to right. Last, in the third condition ('option B'; three replicates), the birds were trained to solve the puzzle box by moving the red side of the sliding door from right to left (Supplementary Video 1). After 4 days of training, all birds were released back into the wild, and three puzzle boxes, with both options available, were installed 250 m apart in each sub-population (Extended Data Fig. 1). We then automatically monitored the individual visits to, and solutions ('solves') at these puzzle boxes, over short-term (20 days of exposure over 4 weeks) and long-term (5 days of exposure, 9 months later) periods.

In the five sub-populations that were seeded with trained demonstrators, knowledge of how to solve the novel puzzle spread rapidly over 20 days of exposure (Fig. 1b). A mean of 75% of the members of each local population (68–83%, $n = 37–96$) solved the puzzle box at least once. The diffusion of this behaviour was clearly sigmoidal (sigmoidal versus linear fit, change in Akaike information criterion (ΔAIC) = 15.31–54.17), except in one replicate (T5, $\Delta AIC = 0.13$). By contrast, many fewer individuals solved the puzzle box in control sub-populations (9–53%, $n = 5–46$; Fig. 1b), in which uptake initially relied on individual innovation. The latency to the first solve, excluding the demonstrator, was significantly longer in control areas than in treatment areas (Welch's two-sample t -test, $t_{(6)} = -16.1$, $P < 0.01$; Fig. 1b), and the total number of solutions was significantly lower ($t_{(6)} = 4.6$, $P = 0.02$; Fig. 1c). There was a striking difference between the replicates that were seeded with alternative solving techniques. In all treatment sub-populations, learning was heavily biased towards the technique that was originally demonstrated ($t_{(8)} = 9.7$, $P < 0.01$; Fig. 1c), while no consistent side bias was observed between the control sub-populations ($t_{(4)} = -0.03$, $P = 0.97$; Fig. 1c).

We derived the social network for each sub-population independently of the social learning experiment, with 10 days' sampling at a grid of

¹Edward Grey Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK. ²Department of Evolution, Ecology and Genetics, Research School of Biology, Australian National University, Canberra, Australian Capital Territory 2600, Australia. ³Department of Anthropology, University of California, Davis, California 95616, USA. ⁴Smithsonian Tropical Research Institute, Ancon 9100, Panama. ⁵Department of Biology, University of Ottawa, Ottawa K1N 9B2, Canada. ⁶Department of Biosciences, Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn TR10 9EZ, UK. ⁷Department of Ecology and Genetics, Uppsala University, 752 36 Uppsala, Sweden.

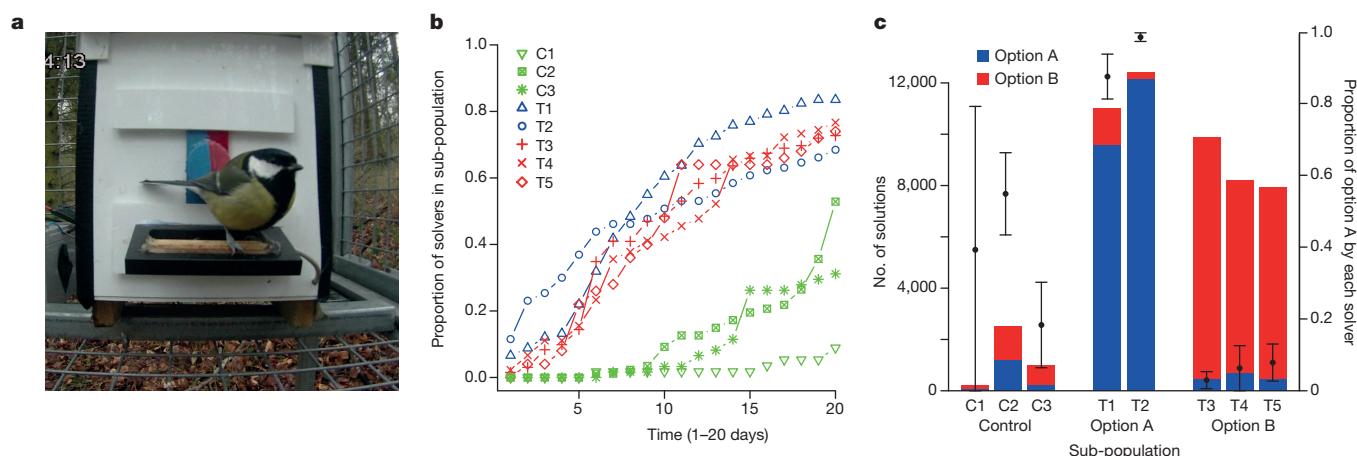


Figure 1 | Cultural diffusion experiment. **a**, A puzzle box in which birds can slide the door open in two directions (from the left, option A; or the right, option B) to access a reward. The puzzle box records the identity, visit duration and solution choice, and it resets after each visit. **b**, Diffusion curves for the treatment sub-populations, with demonstrators (T1–5; $n = 91, 130, 132, 50, 90$, respectively), and the control sub-populations, without demonstrators (C1–3;

$n = 56, 87, 61$, respectively). **c**, The total number of solutions using each option in each replicate (sub-population) (left y axis, shown as stacked bars). The average proportion (dots) of option A performed by individuals, with 95% CI (bars) is shown (right y axis). The total number of solvers was 5, 46 and 19 for C1, C2 and C3 (controls), respectively; 76 and 89 for T1 and T2 (option A), respectively; and 96, 37 and 69 for T3, T4 and T5 (option B), respectively.

sunflower-seed feeders that had been equipped to record visitation data (Extended Data Fig. 2a, b). Co-occurrences (see Methods) were detected using a Gaussian mixture model to isolate clusters of visits in the spatio-temporal data streams²², with repeated foraging associations between individuals forming the basis of social networks (Extended Data Fig. 2b, c). The social networks for all replicates were significantly non-random, even at the most local scale (T1–5, $P < 0.001$), and network-based diffusion analysis was used to quantify the extent to which these social ties predicted the acquisition of behaviour²³. From pooled replicate data, a network diffusion model that included social transmission was overwhelmingly supported over asocial learning ($\Delta\text{AIC} = 1520.7$); the learning rate was estimated to increase by a factor of 12.0 per unit of association with knowledgeable individuals (Extended Data Fig. 3). An effect of age and sex was also supported, with juveniles and males having a faster learning rate (Table 1). These results support a dominant effect of social learning on the emergence of this novel behaviour and also demonstrate that the diffusion of innovation was influenced by the fine-scale patterns of social interactions (Supplementary Video 3).

In all of the experimental replicates, the alternative solution, which was equally difficult and equally rewarded, was performed by at least one individual within the first 6 days of exposure (median, day 4). However, in contrast to most previous studies, in which discovery of an alternative solution led to the progressive erosion of the use of the seeded variant behaviour^{2,5,24}, we observed a pronounced strengthening of traditions

over the rest of the experiment. To analyse this change in behaviour over time, we used a generalized estimating equation model² where the dependent variable was the proportion of solutions using the seeded technique on each day of data collection and the explanatory variables were individuals and replicates. From pooled replicate data, there was strong evidence that the preference for the arbitrary tradition increased over time (coefficient \pm s.e.m. = 0.13 ± 0.02 , $P < 0.001$), with an estimated 14% increase in bias per day (95% confidence interval (CI) = 8–18%; Fig. 2a). This finding is consistent with a conformist transmission bias, with individuals preferentially adopting the more commonly practised variant when solving the puzzle box^{5,7,25,26}. More conclusive evidence for such positive frequency-dependent copying²⁵ was observed when only the first solution for each individual was considered, with birds disproportionately likely to initially adopt the variant used by the majority of their group (sigmoidal versus linear fit, $\Delta\text{AIC} = 38.34$; Fig. 2b).

Individuals thus preferentially learnt the most common option when first learning (Fig. 2b). Yet, remarkably, they also continued to prioritize social information over personal information, matching their behaviour to the common variant even after experiencing an equally rewarding alternative. We analysed the trajectories for those individuals that used both options ($n = 78$). The majority of these individuals (85%, $n = 66$) retained a preference for the seeded variant (for example, see Fig. 2c and Extended Data Fig. 4). Three birds had a strong preference for the uncommon variant, and eight birds switched from the alternative

Table 1 | Network-based diffusion analysis

Network-based diffusion model outputs

Transmission model	ΔAIC (top model)	Σw_i	Social transmission parameter (estimated)	95% CI
Social: multiplicative	0	0.99	12.0	8.8–16.0
T1	–	–	22.4	11.8–30.2
T2	–	–	12.2	8.2–17.1
T3	–	–	7.3	2.9–14.3
T4	–	–	29.8	10.9–42.6
T5	–	–	13.4	8.3–20.02
Social: additive	33.7	0.01	–	–
Asocial	1520.7	0	(Constrained to 0)	–

Individual-level effects

Variable	ΔAIC (top model)	Σw_i	Estimate	Effect size
Age (juvenile or adult)	0	0.99	–0.18	0.70
Sex (F or M)	0	0.97	0.10	1.22
Natal origin (resident or immigrant)	3.9	0.13	0.07	1.16

Summed Akaike weights (Σw_i) and ΔAIC for network-based diffusion models, with maximum-likelihood parameter estimates of social transmission for the five treatment replicates. Estimates and effect sizes are presented for the individual-level variables. The diffusion analyses used a continuous time of acquisition model with a constant baseline learning rate (λ_0), allowing differing social transmission rates in each replicate. F, female; immigrant, dispersed into the study site; M, male; resident, locally born.

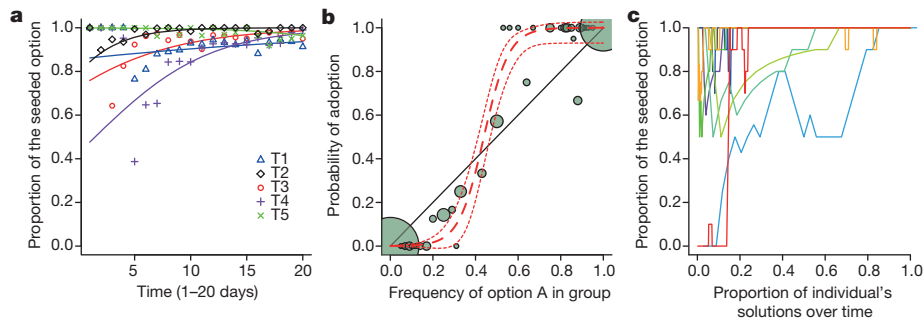


Figure 2 | Evidence of social conformity. **a**, The proportion of solutions using the seeded technique increased significantly over time in each replicate. The points are the proportion of solutions using the seeded technique on each day; the lines show the generalized estimating equation model fit. **b**, Comparison of the frequency of option A in the sub-population with an individual's first learnt option (pooled replicate data from T1–5). The node size represents the

variant to the common variant. However, none of the birds made the reciprocal switch, and only one individual had no significant preference. A subset of birds that dispersed between the experimental replicates (a total of $n=41$, of which 24 were between the two years of the experiment, see below) provided additional evidence. Of 27 birds that moved between replicates with the same seeded tradition, 26 (96%) retained their preference for the common variant. In contrast, of 14 individuals that moved between replicates with different seeded traditions, 10 (71%) changed their behaviour to match the common variant in the new location, 3 retained their initial preference and 1 showed no preference ($\chi^2_{(1)} = 21.6$, $P < 0.001$).

Seeded arbitrary traditions thus formed and persisted in each sub-population (Fig. 2). To investigate the long-term stability of these traditions, we re-installed the puzzle boxes in one replicate of each condition (T1, T3 and C1) over 5 days in the following winter. Substantial turnover in the population had occurred owing to the high mortality rates typical of this species²⁷; on average, only 40% of each sub-population had been present the previous year. No additional demonstrators were trained, and no individual had had contact with the puzzle box in the intervening months. In the control sub-population, all solves ($n = 42$) were performed by only three individuals, all of which had also solved the puzzle box the previous year. However, in the two experimental sub-populations, knowledge of how to solve the puzzle box emerged even

number of individuals ($n = 1–147$). The black line shows the expected result under unbiased copying; the central red line shows the model fit with 95% CI (outer red lines). **c**, The solution trajectories for individuals in the T2 sub-population that used both possible options ($n = 10$). The lines are the running average of the proportion of option A for each individual over the last ten visits, with each colour representing a single individual.

faster than it had the preceding year, both among prior solvers and birds that were inexperienced at the task: in T1, 29 individuals solved the puzzle box a total of 967 times, and in T3, 35 individuals solved the puzzle box a total of 2,329 times (Fig. 3b). The results suggest a strong initial effect of memory, followed by a rapid, oblique transmission facilitated by the greater number of demonstrators than in the initial experiment: on the first day of exposure, 60% (T1) and 82% (T3) of 'solvers' were birds that had solved the puzzle box in the initial experiment, outweighing their representation in the general population (36% in T1 and 46% in T3). The sub-populations also retained their original technique, with the solutions being heavily biased towards the option that had been seeded in the original experiment (Fig. 3b). Intriguingly, among birds that were present in both years, the within-individual bias towards the seeded variant had increased (linear mixed model, $t_{(83)} = 2.80$, $P < 0.01$; Fig. 3c), resulting in arbitrary traditions that were retained and strengthened.

In summary, we show that wild great tits use social learning to acquire novel behaviours and that foraging techniques introduced by few individuals (here only two in each replicate) can spread rapidly to the majority of the population, forming stable arbitrary traditions. Both social network ties and individual characteristics determined the transmission of these foraging techniques²³. The introduced arbitrary traditions were stable over both short-term and long-term periods, becoming increasingly entrenched over two generations. This stability appeared to be a result

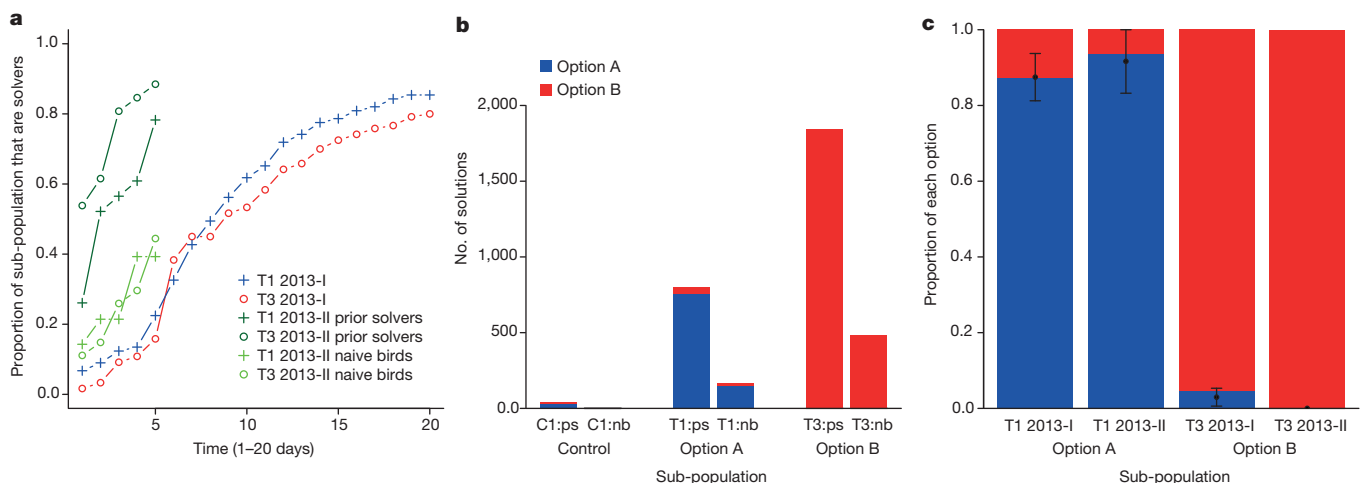


Figure 3 | Local traditions persist across years. **a**, Diffusion curves for the initial exposure (T1 2013-I and T3 2013-I, where I is 20 days) and the second exposure (T1 2013-II and T3 2013-II, where II is 5 days, 9 months later). The cumulative uptake of the behaviour in the second exposure is much higher for prior solvers (T1 2013-II and T3 2013-II, $n = 23$ and 26, respectively) but is also higher for naive birds (T1 2013-II and T3 2013-II, $n = 28$ and 27,

respectively). **b**, The number of solutions using option A or B. In T1, one circuit board failed, so the data are derived from two of three devices. nb, native birds; ps, prior solvers. **c**, The proportion of option A or B used in the initial and second exposure. The histograms show the data for the sub-populations; the dots show the mean proportion of option A performed by individuals, and error bars show the 95% CI.

of informational conformity, with individuals matching their behaviour to the most common variant when first learning and then continuously updating their personal information. Conformity has long been considered a central component of human culture^{25,26,28}, but experimental evidence for its occurrence in wild animals has been limited to a study of food preferences in vervet monkeys⁵. We provide the first experimental demonstration, to our knowledge, of conformist transmission and cultural norms in foraging techniques in a wild animal. Our study argues against the previous view that such behaviour is restricted to the primate lineage^{26,28–30} and calls for a reconsideration of the evolution and ecology of cultural conformity.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 June; accepted 23 October 2014.

Published online 3 December 2014.

- Rogers, E. M. *Diffusion of Innovations* 4th edn (Free Press, 1995).
- Claidière, N., Messer, E. J. E., Hoppitt, W. & Whiten, A. Diffusion dynamics of socially learned foraging techniques in squirrel monkeys. *Curr. Biol.* **23**, 1251–1255 (2013).
- Cantor, M. & Whitehead, H. The interplay between social networks and culture: theoretically and among whales and dolphins. *Phil. Trans. R. Soc. B* **368**, 20120340 (2013).
- Rendell, L. *et al.* Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends Cogn. Sci.* **15**, 68–76 (2011).
- van de Waal, E., Borgeaud, C. & Whiten, A. Potent social learning and conformity shape a wild primate's foraging decisions. *Science* **340**, 483–485 (2013).
- Whiten, A., Hinde, R. A., Laland, K. N. & Stringer, C. B. Culture evolves. *Phil. Trans. R. Soc. B* **366**, 938–948 (2011).
- Whiten, A., Horner, V. & de Waal, F. B. Conformity to cultural norms of tool use in chimpanzees. *Nature* **437**, 737–740 (2005).
- Laland, K. N. & Janik, V. M. The animal cultures debate. *Trends Ecol. Evol.* **21**, 542–547 (2006).
- Warner, R. R. Traditionality of mating-site preferences in a coral reef fish. *Nature* **335**, 719–721 (1988).
- Coussi-Korbel, S. & Frigaszy, D. M. On the relation between social dynamics and social learning. *Anim. Behav.* **50**, 1441–1453 (1995).
- Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B. & Laland, K. N. Identification of the social and cognitive processes underlying human cumulative culture. *Science* **335**, 1114–1118 (2012).
- Laland, K. Social learning strategies. *Learn. Behav.* **32**, 4–14 (2004).
- Galef, B. G. & Laland, K. N. Social learning in animals: empirical studies and theoretical models. *Bioscience* **55**, 489–499 (2005).
- Whiten, A. & Mesoudi, A. Establishing an experimental science of culture: animal social diffusion experiments. *Phil. Trans. R. Soc. B* **363**, 3477–3488 (2008).
- Galef, B. G. in *Oxford Handbook of Comparative Cognition* (eds Zentall, T. R. & Wasserman, E.) Ch. 40, 803–818 (Oxford Univ. Press, 2012).
- Fisher, J. B. & Hinde, R. A. The opening of milk bottles by birds. *Br. Birds* **42**, 347–357 (1949).
- Sherry, D. F. & Galef, B. G. Cultural transmission without imitation: milk bottle opening by birds. *Anim. Behav.* **32**, 937–938 (1984).
- Aplin, L. M., Sheldon, B. & Morand-Ferron, J. Milk-bottles revisited: social learning and individual variation in the blue tit (*Cyanistes caeruleus*). *Anim. Behav.* **85**, 1225–1232 (2013).
- Morand-Ferron, J., Cole, E. F., Rawles, J. E. C. & Quinn, J. L. Who are the innovators? A field experiment with 2 passerine species. *Behav. Ecol.* **22**, 1241–1248 (2011).
- Slagsvold, T. & Wiebe, K. L. Social learning in birds and its role in shaping a foraging niche. *Phil. Trans. R. Soc. B* **366**, 969–977 (2011).
- Aplin, L. M. *et al.* Individual personalities predict social behaviour in wild networks of great tits (*Parus major*). *Ecol. Lett.* **16**, 1365–1372 (2013).
- Psorakis, I., Roberts, S. J., Rezek, I. & Sheldon, B. C. Inferring social network structure in ecological systems from spatio-temporal data streams. *J. R. Soc. Interface* **9**, 3055–3066 (2012).
- Allen, J., Weinrich, M. T., Hoppitt, W. & Rendell, L. Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science* **340**, 485–488 (2013).
- Thornton, A. & Malapert, A. The rise and fall of an arbitrary tradition: an experiment with wild meerkats. *Proc. R. Soc. B* **276**, 1269–1276 (2009).
- Morgan, T. J. H. & Laland, K. The biological bases of conformity. *Front. Neurosci.* **6**, 87 (2012).
- van Leeuwen, E. J. C. & Haun, D. B. M. Conformity in nonhuman primates: fad or fact? *Evol. Hum. Behav.* **34**, 1–7 (2013).
- Bouwhuis, S., Choquet, R., Sheldon, B. C. & Verhulst, S. The forms and fitness cost of senescence: age-specific recapture, survival, reproduction, and reproductive value in a wild bird population. *Am. Nat.* **179**, E15–E27 (2012).
- Haun, D. B. M., Rekers, Y. & Tomasello, M. Majority-biased transmission in chimpanzees and human children, but not orangutans. *Curr. Biol.* **22**, 727–731 (2012).
- de Waal, F. B. Animal conformists. *Science* **340**, 437–438 (2013).
- van Schaik, C. P. Animal culture: chimpanzee conformity? *Curr. Biol.* **22**, R402–R404 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements This project was supported by grants from the BBSRC (BB/L006081/1) and the ERC (AdG 250164) to B.C.S., who was also supported by a visiting professorship at Uppsala University. L.M.A. was also supported by an Australian Postgraduate Award; and A.T., by a BBSRC David Phillips Fellowship (BB/H021817/1). The EGI social networks group, S. Lang and K. McMahon provided assistance in the field, and M. Whitaker produced electronic components for the puzzle boxes.

Author Contributions The study was initially conceived by L.M.A., J.M.-F. and B.C.S., with input from D.R.F., A.C. and A.T. in designing the experiments. Infrastructure to support the work was conceived and developed by B.C.S. The experimental work was led by L.M.A.; and the analysis, by L.M.A. and D.R.F. The manuscript was drafted by L.M.A. and B.C.S., and important contributions were made by all of the other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.M.A. (lucy.aplin@zoo.ox.ac.uk).

METHODS

Study population and area. The study was conducted in a wintering population of tits in Wytham Woods, UK (51° 46' N, 01° 20' W; Extended Data Fig. 1). One thousand and eighteen nest boxes suitable for great tits are installed at this site, with the vast majority of great tits breeding in boxes. Individuals are trapped as nestlings and breeding adults at nest boxes, and are fitted with both a British Trust for Ornithology metal leg ring and a plastic leg ring containing a uniquely identifiable passive integrated transponder (PIT) tag (IB Technology). There is a further mist-netting effort over autumn and winter to tag individuals that immigrate into the population, and we estimate that over 90% of individuals had been PIT-tagged at the time of the study²¹. In this population, great tits form loose fission–fusion flocks of unrelated individuals in autumn and winter. Flocks congregate at patchy food sources and can be observed at bird feeders fitted with PIT-tag-detecting antennae^{21,31}. The experiments were conducted in eight sub-population areas within Wytham Woods that had relatively little short-term between-area movement of individuals (Extended Data Fig. 1). The work was subject to review by the Department of Zoology ethical committee, University of Oxford, and was carried out under Natural England licences 20123075 and 20131205.

Puzzle-box design. The experimental apparatus consisted of an opaque plastic box with a perch positioned in front of a door that could be slid to either side with the bill, to gain access to a feeder concealed behind. Video observations suggested that all great tits used their bill to move the door. The left side of the door was coloured blue, and the right side was coloured red, with a raised front section on the door to allow an easier grip. The concealed feeder contained approximately 500 live mealworms and was refilled up to twice daily. Mealworms are a highly preferred food for great tits (Extended Data Fig. 5), and as live mealworms were used, solvers typically extracted one worm and then carried it away from the puzzle box to kill and eat it (as confirmed by video observations; Supplementary Videos 1 and 2). Each puzzle box was surrounded by a 1 × 1 m² cage with a 5 × 5 cm² mesh that allowed unlimited access by small birds but prevented access by large non-target species such as corvids or squirrels. A freely accessible bird feeder filled with peanut granules was also provided in the cage, at approximately 1 m from the puzzle box. Peanut granules are a much less preferred food source (Extended Data Fig. 5). Each peanut feeder had two access points fitted with RFID antennae and data-logging hardware. This feeder was used to attract the original demonstrator to the location and to record the identity of individuals that did not contact the puzzle box.

All puzzle boxes contained a printed circuit board and motor and were powered by a 12-V sealed battery. The perch also functioned as an RFID antenna that registered the visit duration (the time to nearest second) and the identity of the visiting individual. A 'solve' was recorded if the door was opened during an individual visit to the device, with the side direction also noted. If a solve occurred without an accompanying identified individual, this was recorded as an 'unidentified solve'. One second after the solving bird departed, the door reset itself back to the middle. If more individuals visited before this happened, then a 'scrounge' was recorded, as these individuals were assumed to have taken food from the open door (as confirmed by video observations). The door reset immediately after two individuals were registered scrounging, preventing more than two possible scrounging events per solve (Supplementary Video 2).

Experimental procedure. Two males were captured from each sub-population (11 adults and 5 juveniles) to act as demonstrators. They were captured either by removal from roosting boxes on Sunday night or by mist-netting at a sunflower-seed feeder on Monday morning. They were transferred to individual cages in indoor captive facilities, and over 4 days, each pair of birds was subjected to one of three training regimens using step-wise shaping: (i) given no training and left in the cage with *ad libitum* food (control); (ii) trained to solve the novel puzzle box by pushing the blue side of the door to the right (option B); or (iii) trained to solve the novel puzzle box by pushing the red side of the door to the left (option A). With the exception of 'control' areas, which were clustered in the south of the woodland to avoid cross-contamination, sub-populations were randomly assigned to a training regimen, with both demonstrators from a single sub-population trained on the same technique. During training, the demonstrators were initially exposed to an open puzzle box baited with mealworms, which was then gradually closed over the course of 4 days until the subjects were reliably re-opening it. The other side of the door was fixed during training. On Friday morning, the birds were released back at the site of capture in each respective sub-population. Puzzle boxes for which both options were available and were equally rewarding were installed at three sites 250 m apart on the following Sunday night (Extended Data Fig. 1). These puzzle boxes were run over a 4-week period at each site, continuously operating from Monday to Friday and then removed on Saturday and Sunday, for a total of 20 days of data collection.

Four replicates were conducted in the first year of data collection (December 2012 to February 2013; C1, C2, T1 and T3). At the sites for three of these replicates (C1, T1 and T3), the puzzle boxes were simultaneously re-installed at the same locations for 5 days of further data collection in December 2013. No additional demonstrators

were trained, and no individual had had contact with the puzzle box in the 9 months between the two data collection periods. This second exposure sought to test the long-term stability of social learning at the sub-population level. This study was run before the second year of data collection for the cultural diffusion experiment, to exclude the possibility that dispersing individuals in new replicates could be re-introducing the novel behaviour. An additional four replicates (C3, T2, T4 and T5) were then conducted from December 2013 to February 2014 in new sub-populations, using the same initial protocol.

Data analysis. The local population size for each replicate was defined as comprising all individuals in a replicate that had been recorded at least once at one of the following: the puzzle box, the nearby peanut feeder or the nearest network-logging feeders (operated Saturday and Sunday, see below), during the experimental period (that is, from the weekend following the release of the demonstrators to the weekend after day 20 of operation of the puzzle boxes). When three replicates were compared with the 'persistence' trial in the following year, the local population was defined as all individuals observed at the puzzle box or as all individuals nearby the peanut feeder so that the areas were comparable.

To analyse the results of the initial experiment, we first compared control replicates and treatment replicates, by using Welch's two-sample *t*-tests and by fitting linear and sigmoidal models to the data, with the best model ascertained by the difference in the AIC values³². If individuals were using social information when learning about the puzzle box, then we expected that there would be a difference between the areas seeded with a trained demonstrator (treatment) and those without (control). The replicates were thus compared in terms of latency to first solve (the number of seconds from the beginning of the experimental period, excluding demonstrator) and the total number of solutions. Second, we compared the total number of solutions in the two different experimental treatments. In this case, if a more complex form of social learning than local enhancement to the feeding site was occurring, then we expected a consistent bias towards the seeded variant in the different treatments¹⁴.

To analyse the change in individual and population preferences for option A or B over time, we used a generalized estimating equation (GEE) model² where the dependent variable was the proportion of solutions using the seeded technique on each day of data collection and the explanatory variables were individuals and replicates, weighted by the overall number of solutions per day. The seeded technique (A or B) was initially also included as an explanatory variable but was not significant (coefficient ± s.e.m. = 0.13 ± 0.22, *P* = 0.55). Three individual variables were included in a GEE model: sex, age and natal origin. Sex was determined at capture using plumage coloration; age was determined from breeding records or plumage coloration; and individuals were classed as 'immigrants' if they had dispersed into the study site and 'locally born' if they had been ringed as a nestling in the study site²⁷. Only age was significant (coefficient ± s.e.m. = −0.92 ± 0.20, *P* < 0.001) and was included in the final model (sex, coefficient ± s.e.m. = 0.38 ± 0.22, *P* = 0.08; natal origin, coefficient ± s.e.m. = −0.38 ± 0.22, *P* = 0.08).

If population-level conformity was partly the result of a conformist transmission bias, then at first acquisition we would expect a sigmoidal relationship between the population-level frequency of the option and the probability of adoption, with adoption of the majority option disproportionately more likely than its absolute frequency. By contrast, copying the last individual observed, or random copying, should yield a linear relationship^{25,26}, with the probability of adopting option A or B being roughly equal to its proportion in the overall population. To investigate this, we isolated the first observed solutions by all individuals in all experimental replicates and compared the option choice to the proportion of all previous option A solves observed in the individual's group at that site. The group length was set at 245 s, which was the average group length observed using Gaussian mixture models on temporal patterns of flocking (see below) at network-logging sunflower-seed feeders. Both linear and sigmoidal models were then fitted to the data, with the best model ascertained by the difference in AIC values³².

We further examined the subset of individuals that moved between sub-populations (*n* = 41). This subset included all individuals recorded in more than one experimental replicate, whether within the season (*n* = 17) or between seasons (*n* = 24). No individual was observed in more than two replicates, and this analysis did not include individuals in the 'persistence' trial. A preference for option A or B at each location was defined as more than 75% of all solves for either option A or B in that replicate. Finally, to analyse the change in within-individual bias towards option A or B between the initial experiment and the second-year persistence trials, we used a general linear model where the dependent variable was the number of solves using the seeded variant over the total number of solves for each individual observed in both years. The explanatory variables were treatment type and year, with individual identity as a random effect.

Network data collection and analysis. Sunflower-seed bird-feeding stations had been deployed at 65 locations around Wytham Woods on an approximately 250 × 250 m² grid, as part of long-term research into social network structure in tits^{21,22}. Each

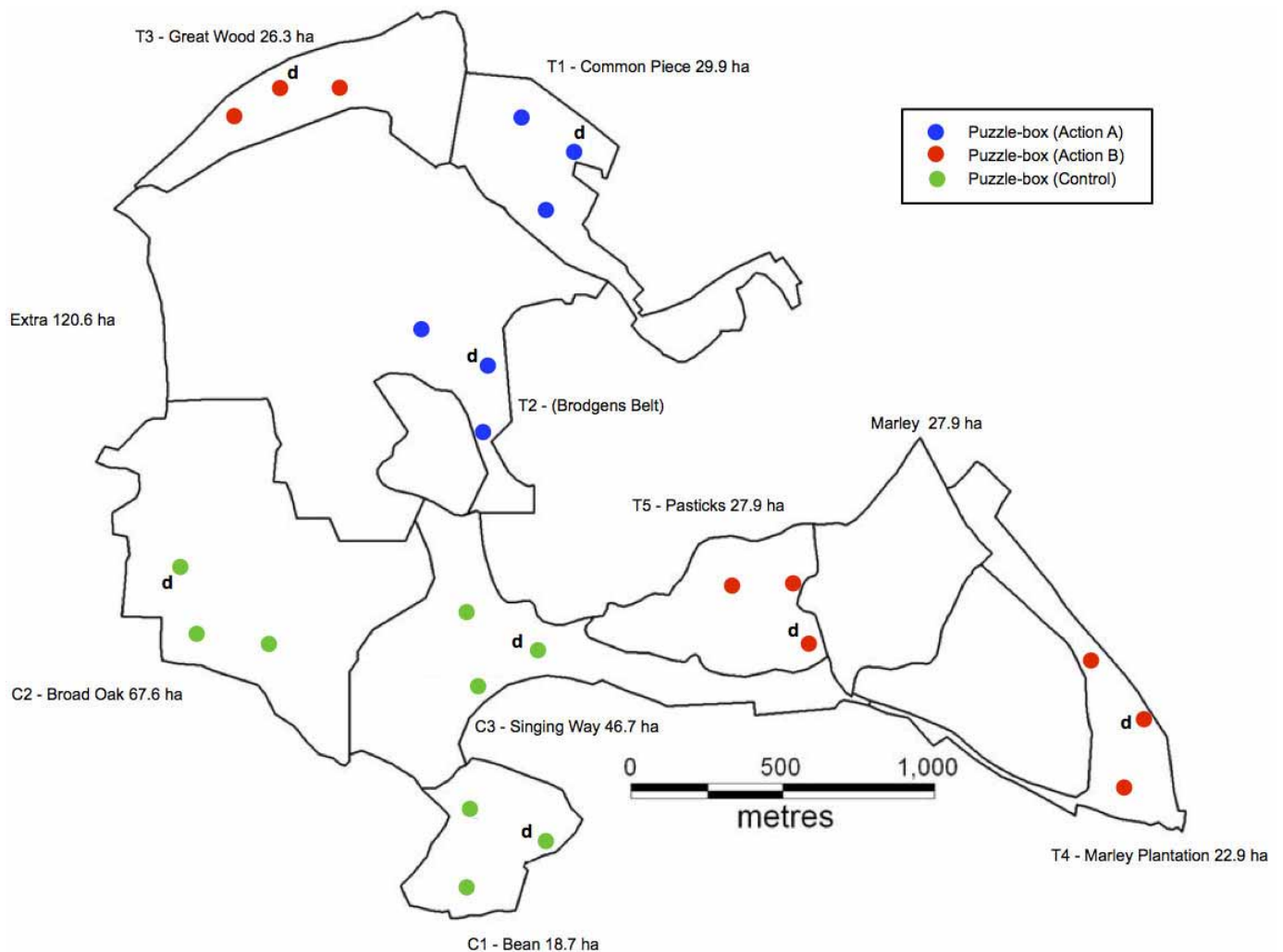
station has two access points, each fitted with RFID antennae and data-logging hardware. The feeding stations automatically opened from dawn to dusk on Saturday and Sunday, scanning for PIT tags every 1/16 s. This study used the data from the eight nearest locations to each set of puzzle boxes, for ten dates within and surrounding the cultural diffusion experiment (the standard logging protocol runs from September to February in Wytham Woods²¹).

Great tits were detected visiting feeding stations and were individually identified by their PIT tags. We then applied a Gaussian mixture model to the spatio-temporal data stream to detect distinct clusters of visits. This method locates high-density periods of feeding activity, thereby isolating flocks of feeding birds without imposing artificial assumptions about group boundaries^{22,33}. A gambit of the group approach³⁴ was used with a simple-ratio index to calculate social associations, where individual association strengths (network edges) were scaled between 0 (never observed foraging together in the same group) to 1 (always observed in the same group and never observed apart). While a single co-occurrence may not be meaningful, our automated data collection method resulted in thousands of repeated group sampling events, allowing social ties between individuals to be built up from multiple observations of co-occurrences over time and across spatial locations. The networks contained 123 (T1), 137 (T2), 154 (T3), 95 (T4) and 110 (T5) nodes; the average edge strengths were 0.09 (T1), 0.05 (T2), 0.08 (T3), 0.07 (T4) and 0.07 (T5). To test whether the networks contained significantly preferred and avoided relationships, we ran permutation tests on the grouping data, controlling for group size and the number of observations, restricting swaps within days and sites^{35,36}. We tested whether the observed patterns of associations were non-random by comparing the coefficient of variance in the observed network with the coefficient of variance in the randomized networks³⁵. The social networks for all replicates were significantly non-random, even at local scales (T1, $P < 0.0001$; T2, $P = 0.0005$; T3, $P < 0.0001$; T4, $P = 0.0002$; and T5, $P = 0.0002$).

Finally, we used network-based approaches to ask whether the behaviour was socially transmitted through foraging associations. Network-based diffusion analysis

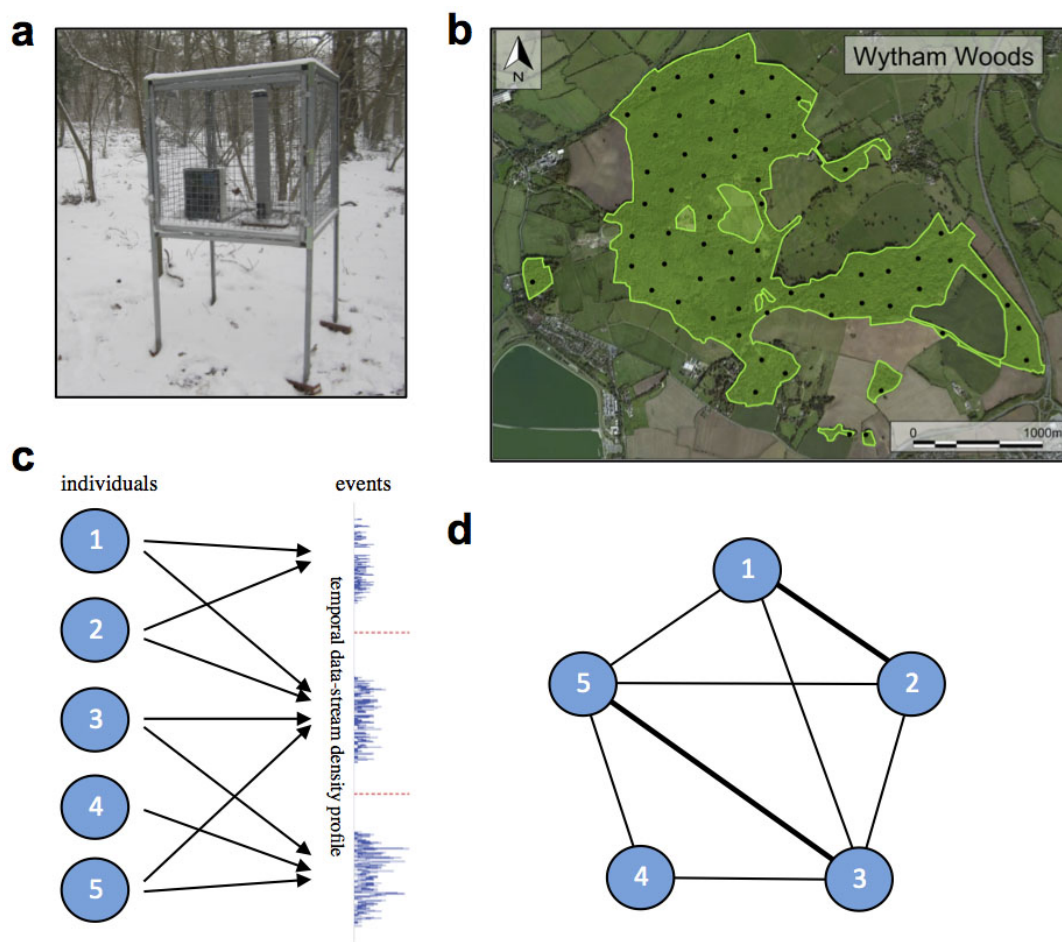
(NBDA) is an approach that tests for social learning by assuming that if social transmission is occurring, then the spread of trait acquisition should follow the patterns of relationships between individuals, with the transmission rate being linearly proportional to the association strength^{23,37,38}. We used NBDA R code v.1.2 (ref. 38), with the time of each individual's first solution (the number of seconds since the beginning of the experiment) entered into the continuous time of acquisition analysis function. Individuals that solved the puzzle box but that did not appear in the social network (that is, that had not been recorded in the standardized weekend logging) were excluded from the analysis. The effects of three individual-level variables were also incorporated into the analysis: sex, age and natal origin. All combinations of NBDA provided in the NBDA R code v1.2 were run, with the social transmission rate allowed to vary for each replicate. An AIC model averaging approach was used to find the best-supported model³⁸.

31. Aplin, L. M., Farine, D. R., Morand-Ferron, J. & Sheldon, B. C. Social networks predict patch discovery in a wild population of songbirds. *Proc. R. Soc. B* **279**, 4199–4205 (2012).
32. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* 2nd edn (Springer, 2002).
33. Farine, D. R. Animal social network inference and permutations for ecologist in R using *asnipe*. *Methods Ecol. Evol.* **4**, 1187–1194 (2013).
34. Franks, D. W., Ruxton, G. D. & James, R. Sampling animal association networks with the gambit of the group. *Behav. Ecol. Sociobiol.* **64**, 493–503 (2010).
35. Whitehead, H. *Analyzing Animal Societies: Quantitative Methods for Vertebrate Social Analysis* (Univ. Chicago Press, 2008).
36. Bejder, L., Fletcher, D. & Brager, S. A method for testing association patterns of social animals. *Anim. Behav.* **56**, 719–725 (1998).
37. Franz, M. & Nunn, C. L. Network-based diffusion analysis: a new method for detecting social learning. *Proc. R. Soc. B* **276**, 1829–1836 (2009).
38. Hoppitt, W., Boogert, N. J. & Laland, K. N. Detecting social transmission in networks. *J. Theor. Biol.* **263**, 544–555 (2010).



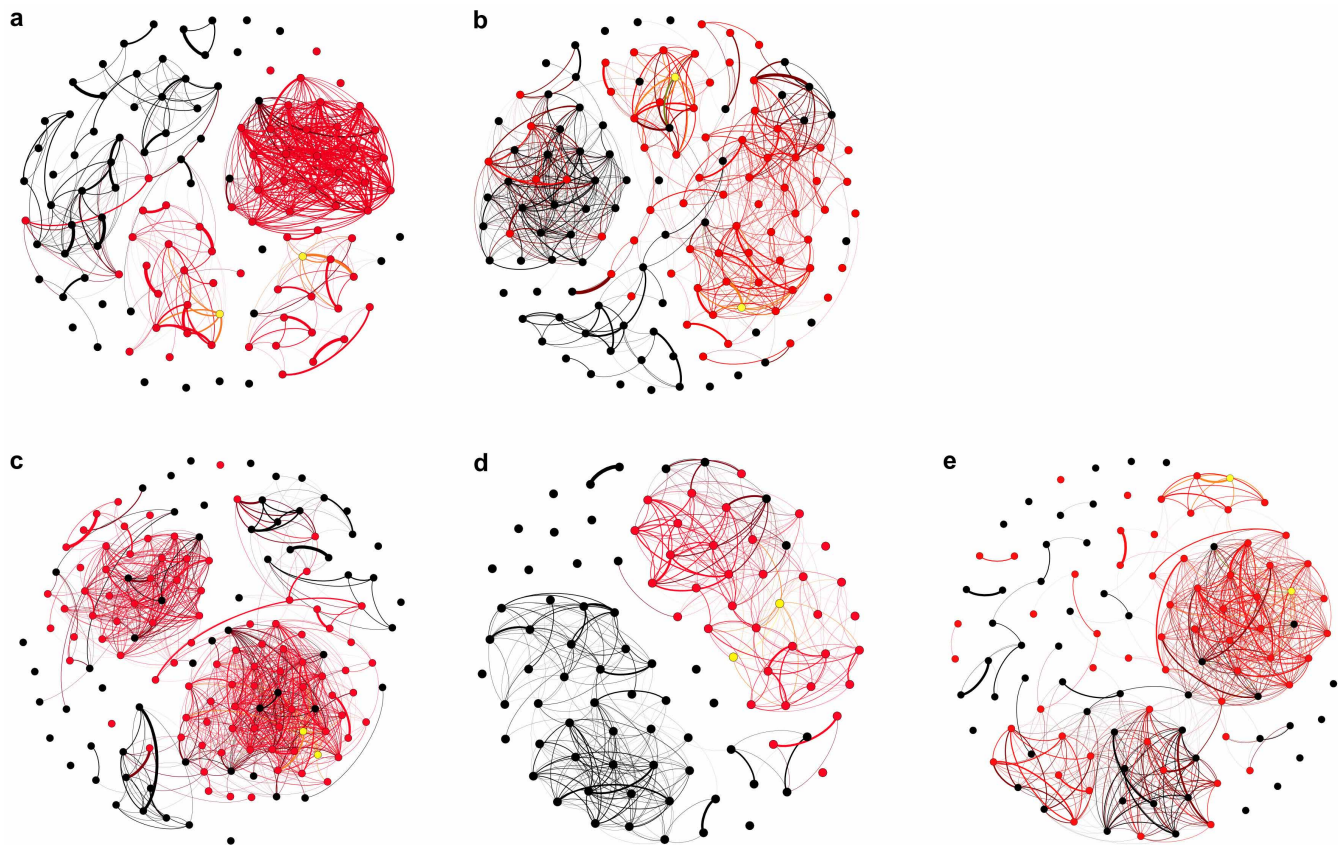
Extended Data Figure 1 | Wytham Woods, UK, ($51^{\circ} 46' \text{ N}$, $01^{\circ} 20' \text{ W}$), showing the location of replicates and puzzle boxes. The total area of Wytham Woods is 385 ha; the location and size of the separate woodland areas within the woods are labelled on the map. The green points indicate the puzzle-box locations for the three control replicates, C1, C2 and C3: Broad Oak, Bean Wood and Singing Way, respectively. The blue points indicate the

puzzle-box locations for the two option A replicates, T1 and T2: Common Piece and Brodgen's Belt, respectively. The red points indicate the puzzle-box locations for the three option B replicates, T3, T4 and T5: Great Wood, Marley Plantation and Pasticks, respectively. (d) indicates the locations where trained demonstrators were caught from and released to.



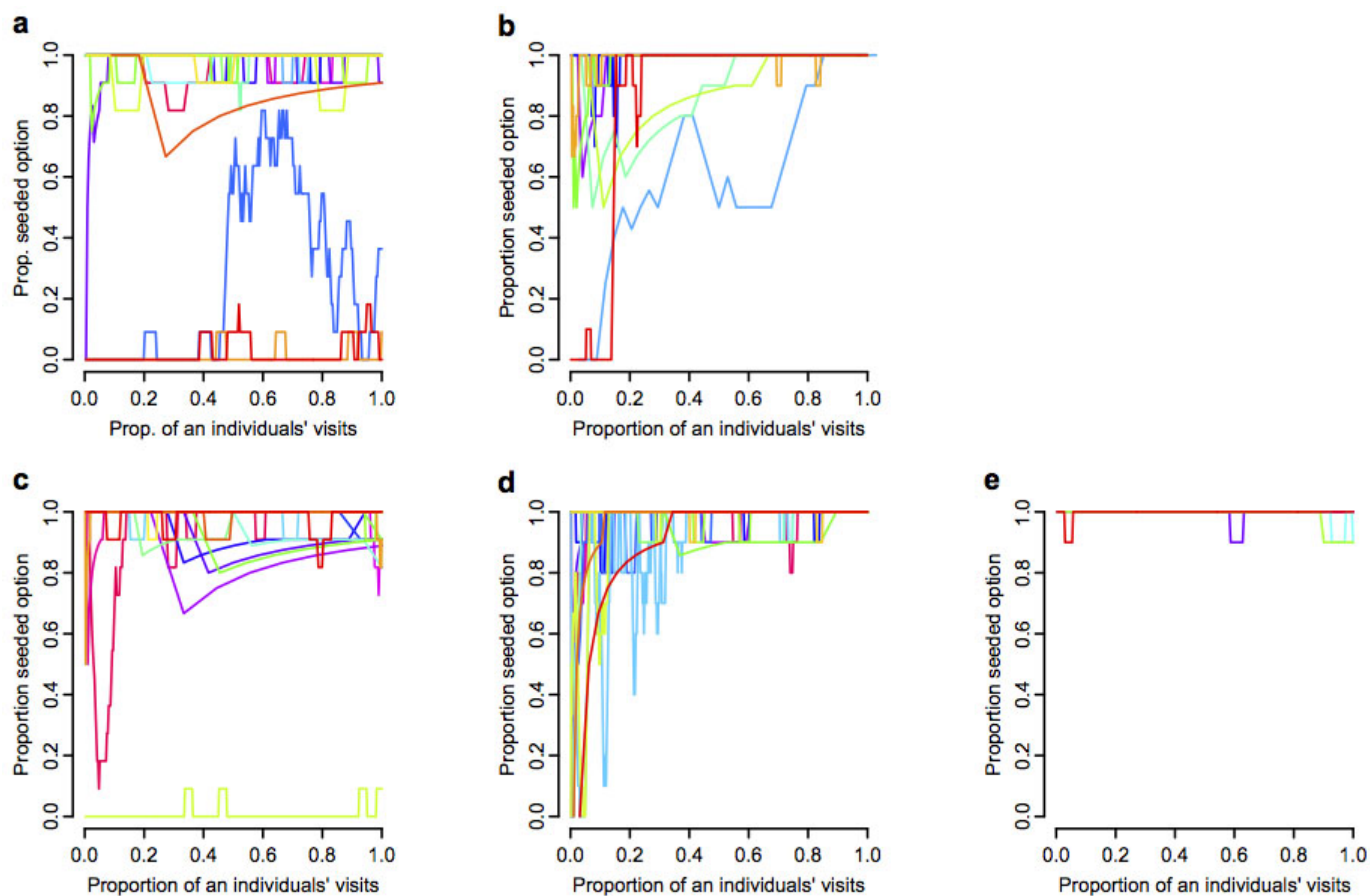
Extended Data Figure 2 | Social network data collection. **a**, Schematic of a feeding station (shut), with sunflower-seed feeder, RFID antennae and data-logging hardware. The cage is to restrict access to small passerines only. **b**, Map of the study area showing the placement of 65 feeding stations. The stations are approximately 250 m apart and open simultaneously from dawn to

dusk on Saturday and Sunday over winter. **c**, Grouping events were inferred from the temporal data stream gained from the feeding stations, with individuals assigned to grouping events in a bipartite network. **d**, Repeated co-occurrences were used to create social networks²².



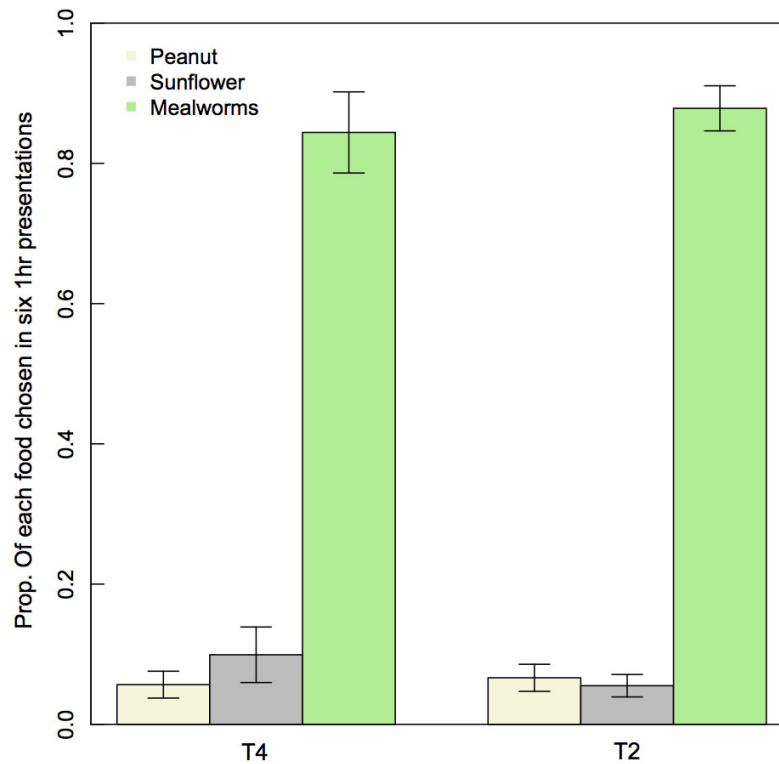
Extended Data Figure 3 | Social networks showing the diffusion of innovation. The red nodes represent individuals that acquired the novel behaviour after 20 days of exposure. The black nodes represent naive individuals. The yellow nodes represent trained demonstrators. The networks are heavily thresholded to show only the links above the average edge strength

for each replicate (T1–T5, 0.09, 0.05, 0.08, 0.07 and 0.07, respectively). **a**, Network for the T1 replicate ($n = 123$). **b**, Network for the T2 replicate ($n = 137$). **c**, Network for the T3 replicate ($n = 154$). **d**, Network for the T4 replicate ($n = 95$). **e**, Network for the T5 replicate ($n = 110$).



Extended Data Figure 4 | Individual trajectories (option A or B) for each replicate. Only individuals that performed both options are included, and individuals that moved between replicates are excluded. The lines are running

proportions of the seeded option for each individual over its last ten visits. **a**, T1 (option A), $n = 30$. **b**, T2 (option A), $n = 10$. **c**, T3 (option B), $n = 19$. **d**, T4 (option B), $n = 15$. **e**, T5 (option B), $n = 4$.



Extended Data Figure 5 | Food preference trials. The birds were presented with a freely available mix of 40 mealworms, 40 peanut granules and 40 sunflower seeds for 1 h on 2 days over 1 week at 6 sites (3 sites for T2 and 3 sites for T4). The trials were conducted 2 weeks after the end of the main experiment,

in March 2014. Food choice was identified from video camera footage, and the trial was halted when all of one prey item was taken. Only great tits were included, but the birds could not be individually identified. The birds clearly preferred the live mealworms to peanut granules or sunflower seeds.

Fundamental properties of unperturbed haematopoiesis from stem cells *in vivo*

Katrin Busch¹, Kay Klapproth^{1*}, Melania Barile^{2*}, Michael Flossdorf^{2*}, Tim Holland-Letz³, Susan M. Schlenner^{4,5}, Michael Reth^{6,7}, Thomas Höfer² & Hans-Reimer Rodewald¹

Haematopoietic stem cells (HSCs) are widely studied by HSC transplantation into immune- and blood-cell-depleted recipients. Single HSCs can rebuild the system after transplantation^{1–5}. Chromosomal marking⁶, viral integration^{7–9} and barcoding^{10–12} of transplanted HSCs suggest that very low numbers of HSCs perpetuate a continuous stream of differentiating cells. However, the numbers of productive HSCs during normal haematopoiesis, and the flux of differentiating progeny remain unknown. Here we devise a mouse model allowing inducible genetic labelling of the most primitive *Tie2*⁺ HSCs in bone marrow, and quantify label progression along haematopoietic development by limiting dilution analysis and data-driven modelling. During maintenance of the haematopoietic system, at least 30% or ~5,000 HSCs are productive in the adult mouse after label induction. However, the time to approach equilibrium between labelled HSCs and their progeny is surprisingly long, a time scale that would exceed the mouse's life. Indeed, we find that adult haematopoiesis is largely sustained by previously designated 'short-term' stem cells downstream of HSCs that nearly fully self-renew, and receive rare but polyclonal HSC input. By contrast, in fetal and early postnatal life, HSCs are rapidly used to establish the immune and blood system. In the adult mouse, 5-fluorouracil-induced leukopenia enhances the output of HSCs and of downstream compartments, thus accelerating haematopoietic flux. Label tracing also identifies a strong lineage bias in adult mice, with several-hundred-fold larger myeloid than lymphoid output, which is only marginally accentuated with age. Finally, we show that transplantation imposes severe constraints on HSC engraftment, consistent with the previously observed oligoclonal HSC activity under these conditions. Thus, we uncover fundamental differences between the normal maintenance of the haematopoietic system, its regulation by challenge, and its re-establishment after transplantation. HSC fate mapping and its linked modelling provide a quantitative framework for studying *in situ* the regulation of haematopoiesis in health and disease.

The paucity of HSCs has largely impeded direct measurements of their functions *in situ*. To determine fundamental properties (frequencies of active HSCs, fluxes between stem and progenitor compartments, residence time and expansion in compartments) of unperturbed steady-state haematopoiesis¹³, we devised an experimental system for inducible genetic marking of HSCs *in situ*. As driver for Cre recombinase (Cre) we used the *Tie2* (also known as *Tek*) locus, which is expressed in embryonic and adult HSCs^{14,15}. We generated a knock-in mutant expressing from the *Tie2* locus a gene encoding codon-improved Cre (iCre) fused to two modified oestrogen receptor binding domains (designated MCM)¹⁶ (Extended Data Fig. 1a–c). We chose this weakly inducible and tightly regulated system to prevent leakiness. The *Tie2*^{MCM} allele was crossed to *Rosa*^{YFP} mice expressing the yellow fluorescent protein (YFP) reporter in a Cre-dependent manner. In the absence of tamoxifen, we did not detect

YFP⁺ haematopoietic cells in bone marrow, thymus and spleen in *Tie2*^{MCM/+} *Rosa*^{YFP} mice ($n = 30$; data not shown). After tamoxifen treatment, MCM becomes active and deletes the stop cassette of the YFP marker gene, thus rendering Cre-expressing cells and their non-Cre-expressing progeny YFP-positive (Extended Data Fig. 1d). Early after treatment (20 days) the labelled cells were almost exclusively HSCs, defined as lineage marker (Lin)[−] Kit⁺ Sca-1⁺ (LSK) CD150⁺ CD48[−] (refs 3, 17 and 18) (Fig. 1a). In a total of 112 adult *Tie2*^{MCM/+} *Rosa*^{YFP} mice, a mean of 1.0% of HSCs were labelled *in situ* after tamoxifen treatment (Extended Data Fig. 1e). Transplantation of a single YFP-marked HSC into genetically conditioned HSC recipients (*Rag2*^{−/−} γ c^{−/−} Kit^{W/W^v}; γ c is also known as *Il2rg*)¹⁹ led to long-term donor HSC engraftment and multilineage reconstitution in primary and secondary recipients (Extended Data Fig. 1f), hence initially labelled cells were functional HSCs (Extended Data Table 1). We ruled out the possibility that HSC numbers and functions were compromised by loss of one *Tie2* allele in *Tie2*^{MCM/+} mice in a series of control experiments (Extended Data Fig. 2).

We estimated by limiting dilution analysis frequencies of HSCs contributing to overall haematopoiesis, and to lymphoid and myeloid lineages (Fig. 1b and Extended Data Fig. 1g–j). At least one out of three HSCs contributed YFP⁺ CD45⁺ progeny in the bone marrow between 6 and 34 weeks after labelling. This is the overall and cumulative frequency regardless of the type of lineage produced, and it represents a lower limit given that all mice with labelled HSCs also contained labelled progeny. Considering 2.8×10^8 total nucleated bone marrow cells per mouse²⁰, and an HSC frequency of 0.006%, a mouse has ~17,000 HSCs; 30% active HSCs indicates that ~5,000 HSCs contributed to normal haematopoiesis within the observation period. In transplantation experiments,

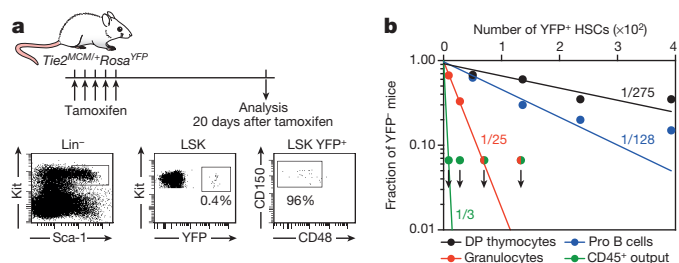


Figure 1 | Inducible HSC labelling in *Tie2*^{MCM/+} *Rosa*^{YFP} mice and frequency estimates on HSC output. **a**, Phenotype of labelled cells 20 days after tamoxifen injection. **b**, Limiting dilution analysis of labelled HSCs and lineage output. The fraction of negative mice for YFP-expressing Lin⁺ CD45⁺ cells (green), granulocytes (red), pro B cells (blue) or double-positive (DP) thymocytes (black) was plotted on a logarithmic scale against the number of YFP⁺ HSCs (Extended Data Fig. 1g–j). Arrows indicate lower detection limit (no negative mice in these groups).

¹Division of Cellular Immunology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany. ²Division of Theoretical Systems Biology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany. ³Division of Biostatistics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany. ⁴Department of Microbiology and Immunology, University of Leuven, B-3000 Leuven, Belgium. ⁵Autoimmune Genetics Laboratory, VIB, B-3000 Leuven, Belgium. ⁶BIOSS, Centre For Biological Signaling Studies, University of Freiburg, Schänzlestraße 18, D-79104 Freiburg, Germany. ⁷Department of Molecular Immunology, Bioll, Faculty of Biology, University of Freiburg, and Max-Planck Institute of Immunobiology and Epigenetics, Stübweg 51, D-79108 Freiburg, Germany.

*These authors contributed equally to this work.

absolute numbers of contributing HSCs are 100-fold lower^{10,11} than our estimate for steady state haematopoiesis. We also determined pathway frequencies from HSCs to granulocytes, and T and B cell progenitors (Fig. 1b and Extended Data Fig. 1g–j). The results revealed correlations between the numbers of labelled HSCs and labelled lineage output, with the (time-averaged) probability for finding labelled granulocytes being five- to tenfold higher than lymphocytes.

To address the fluxes from adult HSCs via stem and progenitor compartments to peripheral lineages, we resolved the output from *in-situ*-labelled HSCs kinetically (Fig. 2a–f). Unexpectedly, up to 3 weeks after induction in adult mice, the label was exclusively retained in HSCs with no label found in LSK CD150⁺CD48⁺ short-term (ST)-HSCs, and LSK CD150⁺CD48⁺ multipotent progenitors (MPPs) (Fig. 2b). Within these downstream stem and progenitor compartments, the first marked cells emerged from 4 weeks onwards (Fig. 2c). Beyond 16 weeks, labelled HSCs had contributed to all analysed progenitor and mature cell lineages, with labelled myeloid cells arising sooner than labelled lymphoid cells (Fig. 2e, f). Analysis of overall bone marrow cells further indicated that the label emanates from marked HSCs (Extended Data Fig. 3a). Very few myeloid progenitors (0.010% common myeloid progenitors (CMPs); 0.001% granulocyte–macrophage progenitors (GMPs); 0.006% megakaryocyte–erythroid progenitors (MEPs)) in the bone marrow were also initially marked (Fig. 2b, asterisk), consistent with weak expression of *Tie2* in myeloid progenitors (<http://www.immgen.org>; data not shown) leading to direct, HSC-independent, labelling. Because of the limited life span of CMPs, GMPs and their progeny (Extended Data Fig. 4), the presence of labelled cells beyond 6 weeks after tamoxifen treatment reflects only cells that have arisen *de novo* from labelled HSCs.

Given the extraordinarily slow label progression out of the adult HSC compartment, we investigated how rapidly HSCs are used during development (Fig. 2g–k and Extended Data Fig. 3b). *Tie2*^{MCM/+} *Rosa*^{YFP} mice were treated *in utero* in midgestation (embryonic day (E) 10.5) with tamoxifen. While initially (E12.5) almost exclusively HSCs (but not erythroid progenitors²¹; Extended Data Fig. 5) were marked in fetal liver, the label progressed within days to progenitors in fetal liver and bone marrow, and by 1 week after birth, equilibrium was nearly reached between labelled HSCs and the entire peripheral system. Hence, HSC use is very rapid (and possibly complete) during development but slow during maintenance of the system (Fig. 2l).

We exploited the kinetic data for adult mice (Fig. 2b–f) to infer the fluxes between stem and progenitor compartments as well as residence time and expansion of the cells in the compartments (Fig. 3). In a given reference compartment (for example, ST-HSCs) the cells lost by onward differentiation (for example, towards MPPs) are replaced by influx from the upstream compartment (for example, HSCs), and by cell production in the reference compartment itself (Fig. 3a). The flux is the product of the rate of differentiation per cell and the total amount of cells undergoing differentiation. The movement of label between compartments contains information on the rate (Fig. 3b), and the ratios of total cell numbers were determined for stem and progenitor compartments (Extended Data Fig. 6a). These considerations form the basis of our model for quantifying the labelling data (Supplementary Methods and Supplementary Discussion).

The steady label frequency of around 1% is consistent with self-renewal of labelled HSCs (Fig. 3c; HSC panel). In compartments downstream from HSCs, labelled cells incrementally replaced non-labelled cells (Fig. 3c). The mathematical model fitted the label frequencies measured up to 240 days after induction, and correctly predicted label frequencies at later time points (Fig. 3c).

We estimated the rates of cell differentiation and net proliferation in the stem and progenitor compartments (Fig. 3d, e, Extended Data Fig. 6b, c and Supplementary Methods). The rate of net proliferation equals the number of cells born per day minus cells lost by death during the same time period. On average, per day, 1 out of 110 HSCs differentiates into an ST-HSC, and 1 out of 22 ST-HSCs differentiates into an MPP (Fig. 3d, f). At the MPP stage, considered a lymphoid–myeloid

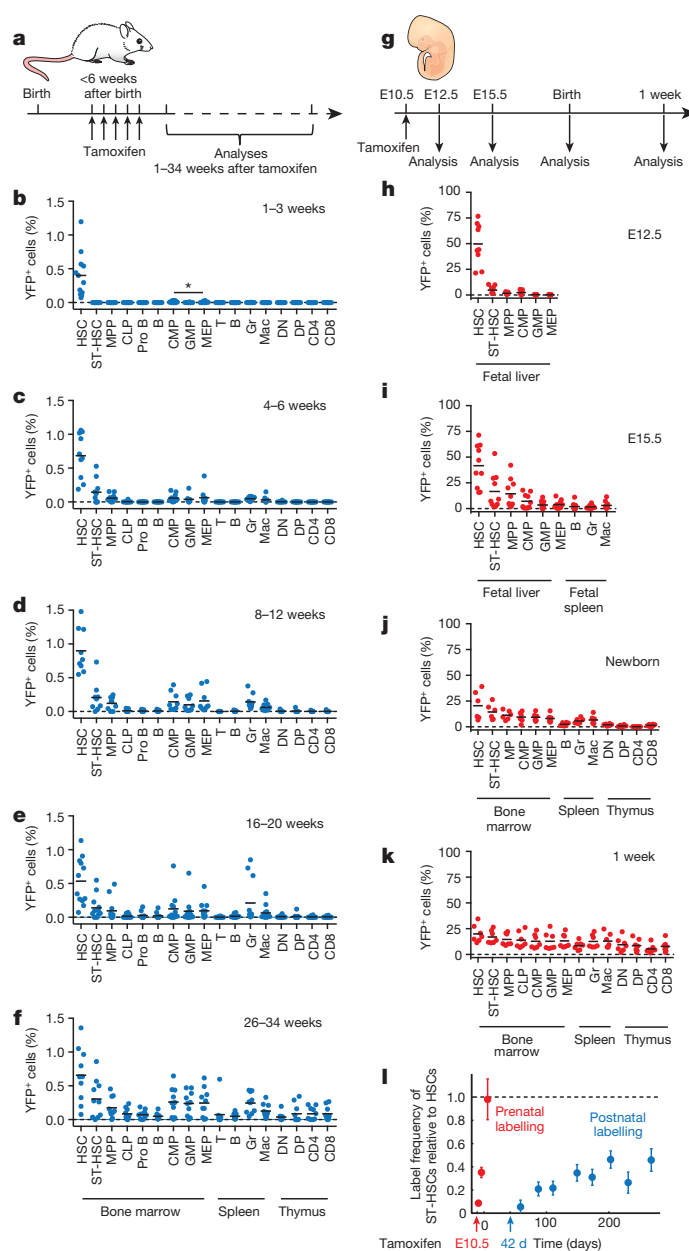


Figure 2 | Label progression through the haematopoietic system during adult maintenance and fetal development. **a**, HSC label induction and output analysis in adult tamoxifen-treated *Tie2*^{MCM/+} *Rosa*^{YFP} mice. **b–f**, Percentages of YFP⁺ cells among the indicated haematopoietic cells in bone marrow, spleen and thymus. Mice were analysed at 1–3 (**b**; $n = 12$), 4–6 (**c**; $n = 10$), 8–12 (**d**; $n = 9$), 16–20 (**e**; $n = 12$) and 26–34 (**f**; $n = 10$) weeks after label induction. Dots represent individual mice, and bars indicate the mean. DN, double negative thymocytes; Gr, granulocytes; Mac, macrophages. **g–k**, *In utero* HSC label induction and analysis in fetal, newborn and 1-week-old *Tie2*^{MCM/+} *Rosa*^{YFP} mice. YFP labelling frequencies were determined at E12.5 (**h**; $n = 9$), E15.5 (**i**; $n = 10$), in newborns (**j**; $n = 6$) and 1 week after birth (**k**; $n = 8$) for the indicated organs and cells. **l**, Kinetic of label progression from HSCs to ST-HSCs in embryonic (red; $n = 66$) and adult (blue; $n = 110$) mice (n per time points; see Supplementary Methods); arrows indicate tamoxifen treatment. Data are mean and s.e.m.

bifurcation point, we estimate that per day 1 out of 46 MPPs generates a CLP, while 1 MPP generates 4 CMPs (Fig. 3d, f). Given that the cell numbers also increase from HSC to ST-HSC and MPP (Fig. 3f), the efflux of cells exceeded influx in all of these compartments. To maintain compartment size, this flux difference is balanced by net proliferation (efflux minus influx = net proliferation). The rates of net proliferation

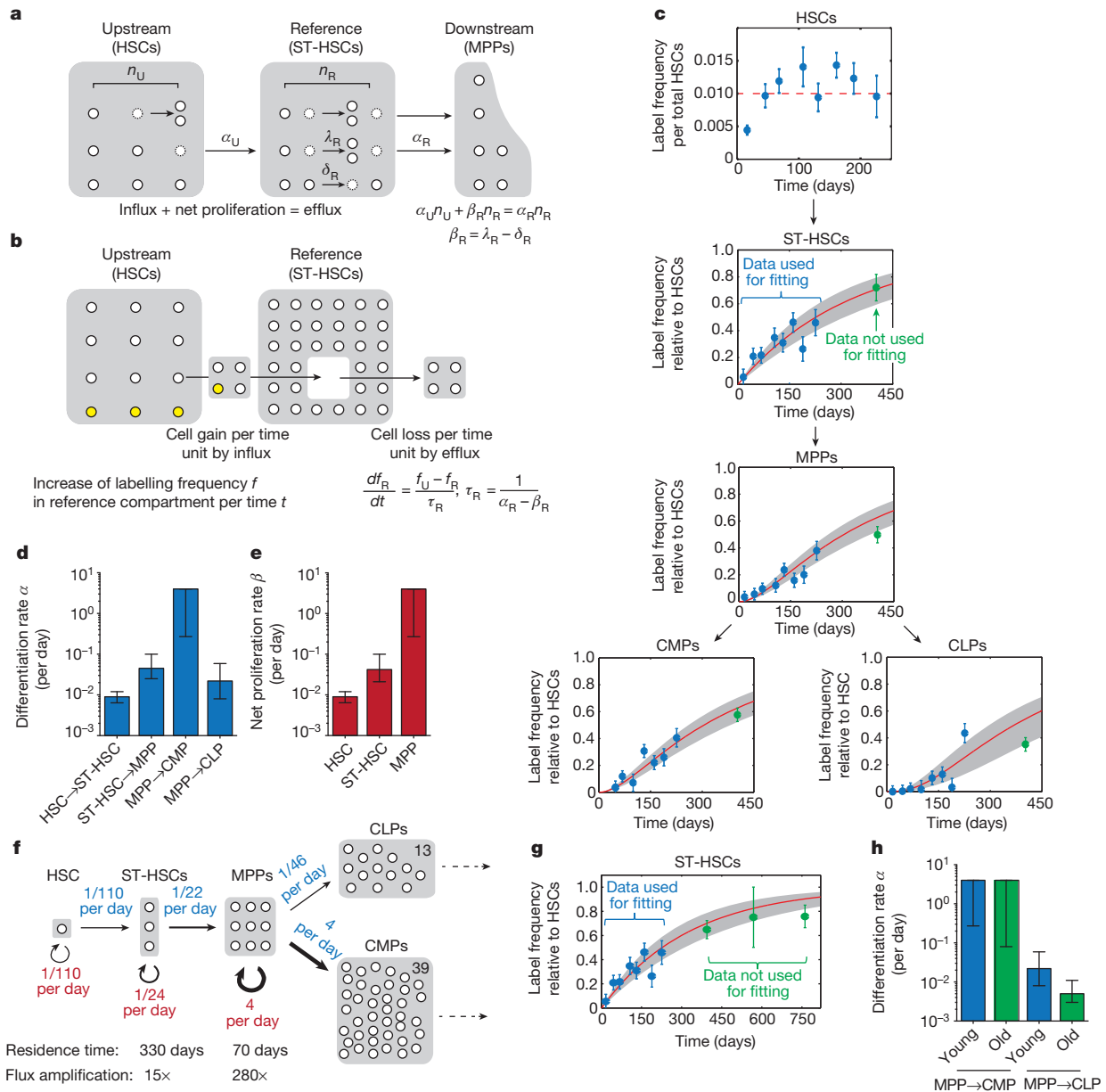


Figure 3 | Inference of stem and progenitor cell differentiation, and proliferation from label progression. **a**, At steady state, the rate of cell loss in a reference compartment due to cell differentiation and death is balanced by cell influx from the upstream compartment, and by proliferation within the reference compartment. This balance relates the total upstream and reference compartment sizes (n_U and n_R , respectively) to the rates of cell differentiation (α_U and α_R) and to net proliferation (proliferation – death) ($\beta_R = \lambda_R - \delta_R$). **b**, The label frequency in the reference compartment (f_R) equilibrates over time with the label frequency in the upstream compartment (f_U). The time for label equilibration $\tau_R = 1/(\alpha_R - \beta_R)$ (residence time) is determined by how rapidly cells are lost from the reference compartment. **c**, HSC label over time (blue dots; red dashed line, average), label progression (blue dots, with s.e.m.,

increased from HSC via ST-HSC to MPP compartments, in parallel with the differentiation rates (Fig. 3d, e).

Together, the differentiation rate and net proliferation determine how close a compartment operates to self-renewal. To quantify the degree of self-renewal, we define the residence time in a compartment as the time period in which the compartment would decay to 37% ($1/e$) of its size, if all influx were switched off. The residence time is determined by the duration a cell and its progeny spend in the compartment before being lost by differentiation or cell death. The compartment residence times can be estimated from the labelling data (Supplementary Methods).

as in Fig. 2l), and mathematical model fitting (red lines, best fit; grey shades, 95% confidence bands). Data measured at ~400 days were not used for fitting (green points, $n = 11$ mice) but have been predicted correctly by the model. **d, e**, Inferred average rates of cell differentiation (**d**) and net proliferation (**e**) for stem and progenitor cell compartments (with 95% confidence intervals). **f**, Dynamics of stem and progenitor cell compartments inferred from the experimental data. Relative compartment sizes are symbolized by grey boxes, and magnitude of fluxes by arrow width. **g**, The model based on data in young mice (blue dots) also predicts data obtained in very old mice (green dots). **h**, Inferred rates of myeloid and lymphoid differentiation from MPPs in younger (7–238 days after label induction; CMPs only considered beyond 42 days) ($n = 110$) and older (332–802 days) ($n = 41$) mice.

For HSCs, given that the labelling frequencies are maintained over time (Fig. 3c) despite efflux (Fig. 3d), net proliferation ensures complete self-renewal, and the residence time is theoretically infinite. Linking our estimate of the HSC net proliferation rate (~1 out of 110 per day; Fig. 3e, f) to proliferation measurements would imply that in assays using 5-bromo-2'-deoxyuridine (BrdU) ~1% of HSCs were labelled per day if HSCs lived indefinitely (and ~2% if HSC lifetime was 100 days). This figure is in the order of magnitude of the reported ~4% BrdU labelling per day in the 'HSC-1' population (as a subset of LSK SLAM-defined HSCs)¹⁸, suggesting that the Tie2⁺ HSCs we label *in situ* reside at the top of

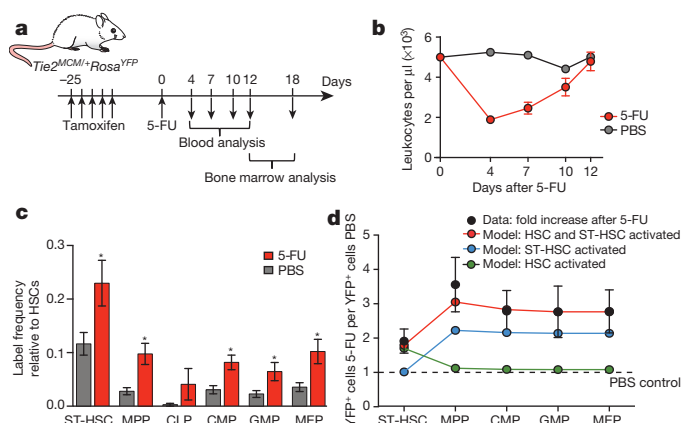


Figure 4 | *In situ* HSC response to 5-FU challenge. **a**, Experimental outline. **b**, Mean of absolute leukocyte numbers in peripheral blood on the indicated days after 5-FU ($n = 12$, red dots) or PBS ($n = 9$, black dots) injections. **c**, Labelling frequencies of the indicated populations relative to HSCs after 5-FU ($n = 18$, red bars) or PBS ($n = 15$, grey bars) (mean from days 12 and 18). * $P = 0.031$ (ST-HSCs), 0.003 (MPPs), 0.213 (CLPs), 0.002 (CMPs), 0.034 (GMPs) and 0.013 (MEPs) (two-tailed t -test assuming non-equal variances; 5-FU versus PBS). **d**, Simulated increase (coloured dots and lines) in YFP labelling frequencies considering participation of HSCs and/or ST-HSCs versus experimental data (black dots; ratios of 5-FU over PBS frequencies, taken from **c**). Error bars denote s.e.m.

the HSC hierarchy. For ST-HSCs, net proliferation almost accounts for the efflux towards MPPs, and the compartment requires only minimal influx from HSCs. Hence, even ST-HSCs operate near self-renewal, their residence time is exceedingly long (330 days), and label progression does not reach equilibrium within the ~ 2 -year lifetime of a mouse (Fig. 3g). Substantial self-renewal (residence time 70 days) was found even at the MPP stage. Proliferation of MPPs leads to an efflux of cells into common lymphoid progenitors (CLPs) and CMPs that is approximately ~ 280 times the influx from ST-HSCs (Fig. 3f), making the MPPs a key amplifier. We also analysed progenitor compartments downstream from CLPs (pro B) and CMPs (MEPs and GMPs) (Extended Data Fig. 6d).

An imbalance between myeloid and lymphoid production has been viewed as an age-dependent HSC property²². We used fate mapping to re-address this question independent of transplantation, and found a marked myeloid bias that, however, was only marginally accentuated with age through relative loss of lymphoid potential (Fig. 3h). Such bias could be caused by skewed differentiation from a common progenitor, or by preferential proliferation in the myeloid branch. Phenotypes and designations of tested progenitors at putative branch points are shown in Extended Data Fig. 7. In both scenarios, which are not mutually exclusive, production of CMPs was several-hundred-fold larger than that of CLP.

To examine the responsiveness of haematopoiesis to perturbation, we challenged HSC-labelled mice with a single injection of 5-fluoruracil (5-FU), a cytotoxic agent causing a transient leukopenia in the blood (Fig. 4a, b). After peripheral rebound, we observed higher stem and progenitor cell labelling frequencies, relative to HSC labelling, than in untreated mice (Fig. 4c). This accelerated label equilibration between HSCs and subsequent compartments after haematopoietic injury indicates feedback inhibition on HSCs output under steady state. These data are fit by a model in which the kinetics of net proliferation and differentiation are accelerated (at least) in both HSC and ST-HSC compartments, but not in either alone (Fig. 4d and Supplementary Discussion).

Output from highly polyclonal HSCs in adult haematopoiesis *in situ* (Fig. 1b) is in contrast to oligoclonal HSC activity found in transplantation experiments^{7–12}. To address this discrepancy, we followed the fate of *in-situ*-marked HSCs after conventional bone marrow transplantation. Adult *Tie2^{MCM1/+}Rosa^{YFP}* mice were treated with tamoxifen, and Lin[−]Kit⁺ bone marrow cells from each donor (CD45.2) were transplanted into

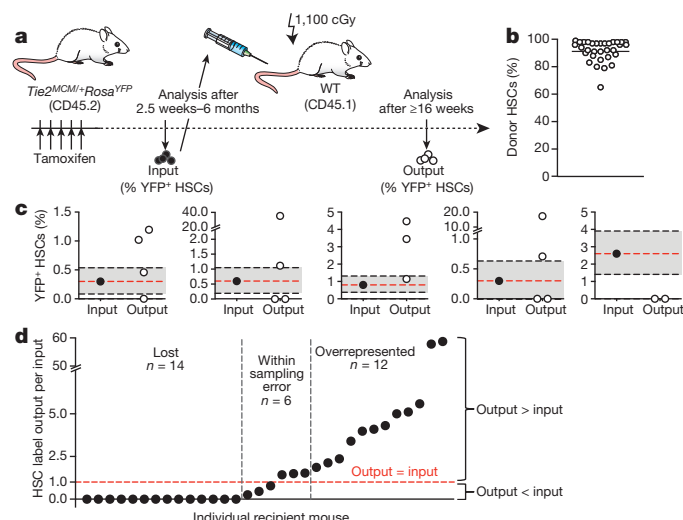


Figure 5 | Fate of *in-situ*-labelled HSCs after bone marrow transplantation. **a**, Experimental outline. **b**, Chimaerism of donor HSCs in recipient mice ($n = 32$) (mean, horizontal line). **c**, Frequencies of YFP⁺ HSCs (black dots; level indicated by red dashed lines) in individual donor mice, and of YFP⁺ donor HSCs (white dots) in the corresponding recipient mice. Representative data from a total of 11 donors and 32 recipients are shown. Grey shaded areas denote 95% confidence intervals of sampling errors. **d**, Ratio of engrafted HSC label output over donor HSC input. Red dashed line (ratio = 1) indicates equal output and input ($n = 32$ recipients).

1–4 lethally irradiated recipients (CD45.1) (Fig. 5a). For each donor, the HSC labelling frequency was recorded before transplantation ('input'). Despite uniformly strong donor HSC engraftment after 16–18 weeks (average 90%; Fig. 5b), the percentages of YFP-marked cells among total donor HSC ('output') were highly variable compared to the input frequencies in individual recipients (Fig. 5c). Input and output were roughly equal (within sampling error) in only 6 out of 32 recipients, whereas in most recipients donor HSCs were either lost (14 out of 32) or overrepresented (12 out of 32) (Fig. 5d). We estimate that on average 1 out of 33 donor HSCs engrafted (Extended Data Fig. 8a–c). In two extreme cases, YFP⁺ donor HSCs represented only 0.3% or 0.6% of the input, but 17% or 35%, respectively, of the output, suggesting much stronger proliferation than for HSCs under steady state (Extended Data Fig. 8d and Supplementary Discussion). In summary, the participation of individual engrafted HSCs to the repopulation of the bone marrow is highly heterogeneous.

Inducible labelling of HSCs in normal mouse bone marrow showed that during development HSCs are rapidly used to establish the haematopoietic system. Once this is accomplished, individual HSCs are only rarely active, but over time a large portion of HSCs contributes to adult haematopoiesis. Indeed, although the mean HSC labelling frequency was low, all mice with marked HSCs produced labelled progeny, indicating that a large fraction, or at least 30%, of all HSCs contributes to haematopoiesis in adult mice after label induction. By contrast, previous work based on barcoding showed that few HSCs actively drive haematopoiesis after transplantation. We re-addressed HSC diversity in the wake of transplantation, avoiding potential pitfalls of cellular heterogeneity of mixing experiments. The observed HSC oligoclonality is hence a hallmark of post-transplantation but not normal unperturbed haematopoiesis. These findings indicate that experimental and possibly also clinical HSC transplantations are based on a much smaller stem cell foundation than physiological haematopoiesis.

HSC proliferation has often been taken as a proxy for asymmetric cell division (for example, ref. 23) and, also indirectly, as measure of differentiation rates. Proliferation may, however, not only yield differentiating progeny but also compensate for cell loss, precluding proliferation as an unambiguous marker of differentiation rates. Here, we quantified haematopoietic flux based on label progression from HSCs. Our data

support in principle an order of differentiation *in situ* from HSCs to ST-HSCs and MPPs and onwards. However, in divergence from the idea of a continuous stream from the top of a haematopoietic pyramid, a very low flux emanated from HSCs. While ST-HSCs are relatively short-lived on transplantation, *in situ* this compartment is exceedingly long-lived because ST-HSC self-renewal is almost sufficient to make up for the cell loss by differentiation. We hence identified this compartment as the primary source of haematopoietic maintenance in mice. This reservoir property of ST-HSCs readily explains the apparent HSC-independence of haematopoiesis noted in a recent report²⁴. However, to maintain the ST-HSC compartment in the long run (>1 year), it requires continuous input from HSCs; we estimate that per day 150 HSCs feed into this compartment (17,000 total HSC \times 1/110 differentiating per day). Hence, true HSC deficiency may go unnoticed for extended periods of time while functionally impaired ST-HSC and MPP compartments would cause rapid signs of acute bone marrow failure (Extended Data Fig. 9).

Collectively, HSCs act in development as founding stem cells, and in adult mice as replenishing cells, ST-HSCs as long-term amplifying cells, and MPPs as intermediate-term amplifying cells. The described fate-mapping system may also visualize responses to haematopoietic challenges imposed by cancer, infections, cachexia or ageing. The accelerated HSC output in response to haematopoietic injury by treatment with 5-FU underscores this outlook.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 July 2014; accepted 19 January 2015.

Published online 11 February 2015.

- Smith, L. G., Weissman, I. L. & Heimfeld, S. Clonal analysis of hematopoietic stem-cell differentiation *in vivo*. *Proc. Natl Acad. Sci. USA* **88**, 2788–2792 (1991).
- Osawa, M., Hanada, K., Hamada, H. & Nakauchi, H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* **273**, 242–245 (1996).
- Kiel, M. J., Yilmaz, O. H., Iwashita, T., Terhorst, C. & Morrison, S. J. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109–1121 (2005).
- Sieburg, H. B. *et al.* The hematopoietic stem compartment consists of a limited number of discrete stem cell subsets. *Blood* **107**, 2311–2316 (2006).
- Dykstra, B. *et al.* Long-term propagation of distinct hematopoietic differentiation programs *in vivo*. *Cell Stem Cell* **1**, 218–229 (2007).
- Abramson, S., Miller, R. G. & Phillips, R. A. The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems. *J. Exp. Med.* **145**, 1567–1579 (1977).
- Keller, G., Paige, C., Gilboa, E. & Wagner, E. F. Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haematopoietic precursors. *Nature* **318**, 149–154 (1985).
- Lemischka, I. R., Raulet, D. H. & Mulligan, R. C. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* **45**, 917–927 (1986).
- Dick, J. E., Magli, M. C., Huszar, D., Phillips, R. A. & Bernstein, A. Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of *W/W^u* mice. *Cell* **42**, 71–79 (1985).
- Gerrits, A. *et al.* Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood* **115**, 2610–2618 (2010).
- Lu, R., Neff, N. F., Quake, S. R. & Weissman, I. L. Tracking single hematopoietic stem cells *in vivo* using high-throughput sequencing in conjunction with viral genetic barcoding. *Nature Biotechnol.* **29**, 928–933 (2011).
- Naik, S. H. *et al.* Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229–232 (2013).
- Harrison, D. E., Lerner, C., Hoppe, P. C., Carlson, G. A. & Alling, D. Large numbers of primitive stem cells are active simultaneously in aggregated embryo chimeric mice. *Blood* **69**, 773–777 (1987).
- Yano, M. *et al.* Expression and function of murine receptor tyrosine kinases, TIE and TEK, in hematopoietic stem cells. *Blood* **89**, 4317–4326 (1997).
- Hsu, H. C. *et al.* Hematopoietic stem cells express Tie-2 receptor in the murine fetal liver. *Blood* **96**, 3757–3762 (2000).
- Zhang, Y. *et al.* Inducible site-directed recombination in mouse embryonic stem cells. *Nucleic Acids Res.* **24**, 543–548 (1996).
- Weissman, I. L. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* **100**, 157–168 (2000).
- Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, 102–116 (2013).
- Waskow, C. *et al.* Hematopoietic stem cell transplantation without irradiation. *Nature Methods* **6**, 267–269 (2009).
- Boggs, D. R. The total marrow mass of the mouse: a simplified method of measurement. *Am. J. Hematol.* **16**, 277–286 (1984).
- Gomez Perdiguero, E. *et al.* Tissue-resident macrophages originate from yolk sac-derived erythro-myeloid progenitors. *Nature* <http://dx.doi.org/10.1038/nature13989> (2014).
- Geiger, H., de Haan, G. & Florian, M. C. The ageing haematopoietic stem cell compartment. *Nature Rev. Immunol.* **13**, 376–389 (2013).
- Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
- Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322–327 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Blum for blastocyst injections, T. Arnspurger, N. Maltry and S. Schäfer for technical assistance, and the animal facility at the DKFZ for expert mouse husbandry. We are grateful to N. Dietlein, H. J. Fehling, T. Feyerabend, C. Opitz, J. Rodewald, A. Roers, K. Rajewsky and A. Trumpp for discussions. K.B. was initially funded through the International Graduate School in Molecular Medicine, Ulm. M.R. was supported by the DFG through EXC294, TRR130 and SFB746. T.H. is a member of CellNetworks, and was supported by BMBF e:Bio grant T-Sys (Fkz 031614) and EU-FP7 Marie Curie ITN Quantitative T cell immunology (QuanTI). H.-R.R. was supported by DFG-SFB 873-project B11, DFG-SFB 938 project L, ERC Advanced Grant no. 233074, and the Helmholtz Alliance on preclinical cancer models (PCCC).

Author Contributions K.B. designed and performed experiments, K.K. and S.M.S. performed experiments, M.B. and M.F. performed theoretical modelling, T.H.-L. made initial bioinformatic analyses, M.R. contributed to the tamoxifen-regulated Cre construct, T.H. developed mathematical models and wrote the paper, and H.-R.R. conceived and supervised the study, and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.-R.R. (hr.rodewald@dkfz.de) or T.H. (t.hoefer@dkfz.de).

METHODS

Generation of *Tie2*^{MCM/+} knock-in mice. For inducible marking of HSC *in situ* we used the *Tie2* locus as driver for Cre recombinase (Cre). Because conventional Cre was insufficient for inducible HSC labelling in our first *Tie2* knock-in mutant, we constructed a second one in which we inserted a gene encoding codon-improved Cre (iCre) fused to two modified oestrogen receptor binding domains, referred to as MCM⁺, into the first exon of the *Tie2* locus by homologous recombination in embryonic stem cells. The targeting strategy is depicted schematically in Extended Data Fig. 1a. The targeting construct consisted, from 5' to 3', of a short homologous arm (from -1898 to -1 base pairs (bp)) upstream of the ATG start codon of the *Tie2* gene (ENSMUSG0000006386), a splice donor site, an intron and a splice acceptor site (intron), all taken from the rabbit β -hemoglobin (*HBB2*) gene²⁵, the coding sequence of codon-improved Cre²⁶ flanked by two mutated oestrogen receptor sites^{16,27}, a polyadenylation signal (pA) from the rabbit β -hemoglobin gene, a FRT-flanked neomycin (*Neo*) resistance gene, a long homologous arm (nucleotides +4 to +4983 of the *Tie2* gene, considering the adenine in the ATG start codon as position +1), and finally the diphtheria toxin subunit A gene (*DT-A*) for selection against random integration. This weakly inducible and tightly regulated system was chosen to prevent leakiness²⁸. The complete nucleotide sequence of the final targeting vector can be obtained from the authors on request. Gene targeting experiments were performed in E14.1 embryonic stem cells. Correct homologous recombination of the targeting vector resulted in replacement of the start codon (nucleotides +1 to +3) in the first exon of the *Tie2* locus with the MCM cassette. Correctly targeted clones were first identified by PCR. Targeted embryonic stem cell clones were transiently transfected with a plasmid (pCAGGS-FlpE-GFP) expressing Flp-recombinase (gift from H. J. Fehling) to delete the *Neo* cassette. Site-specific integration and *Neo* deletion were confirmed by Southern blotting. DNA was digested with BspHI. Blots were hybridized with a radiolabelled 1.0-kb iCre-specific probe. All *Tie2* gene sequences upstream and downstream of exon 1 are preserved in the *Tie2*^{MCM} allele. A *Neo*-deficient embryonic stem cell subclone (*Tie2*^{MCM/Neo}) was injected into C57BL/6 blastocysts, and chimaeric mice were backcrossed to C57BL/6 mice to transmit the *Tie2*^{MCM} allele. Heterozygous *Tie2*^{MCM/+} mice are fertile and show no apparent abnormalities. Homozygous *Tie2*^{MCM/MCM} embryos die between E9.5 and E12.5, as previously described for *Tie2*^{-/-} mice^{29,30}.

Induction of reporter gene expression by tamoxifen. *Tie2*^{MCM/+} mice were crossed to *Rosa*^{YFP} (Gt(ROSA)26Sor^{tm1.1(EYFP)Cos}) mice³¹. Tamoxifen (1 g; Sigma T5648) was dissolved in 4 ml ethanol absolute and 36 ml peanut oil (Sigma P2144) at 55 °C. Aliquots of tamoxifen (25 mg ml⁻¹) were stored at -20 °C and heated to 37 °C shortly before usage. Mice were injected daily on 5 consecutive days with 1 mg tamoxifen intraperitoneally. For *in utero* tamoxifen treatments the day of the plug was regarded as day 0.5 of the pregnancy (E0.5). Pregnant mice were treated by oral gavage on E7.5 or E10.5 with a single dose of 2.5 mg tamoxifen and 1.75 mg progesterone (Sigma P0130) to counteract late fetal abortions. Delivery of the pups was routinely assisted by caesarean section at E20.5, and mice were raised by foster mothers.

Flow cytometry. Bone marrow cells were flushed from femurs, tibiae, coxa and humeri using PBS supplemented with 5% heat-inactivated FCS in PBS. Cells were filtered through a 20- μ m filter (Falcon). Spleens, thymi and fetal organs were directly mashed in a 20- μ m filter with a plunger of a syringe. Fc receptors were blocked by incubating cells in 5% FCS with purified mouse IgG (500 μ g ml⁻¹, Jackson ImmunoResearch Laboratories). All stainings were performed in 5% FCS on ice for 30 min with optimal dilutions of commercially-prepared antibodies. Reagents used were CD3e allophycocyanin (APC) (17A2), CD3e eFluor780 (17A2), CD3e phycoerythrin (PE) (145-C11), CD4 PE-Cy7 (GK1.5), CD8 APC (53-6.7), CD11b PE (M1/70), CD11b PE-Cy7 (M1/70), CD11b PerCP Cy5.5 (M1/70), CD16/32 PE-Cy7 (93), CD16/32 PE-Cy5.5 (93), CD19 PerCP-Cy5.5 (ID3), CD34 eFluor660 (RAM34), CD45 PE-Cy7 (30-F11), CD45.1 eFluor660 (A20), CD45.2 PE (104), CD48 APC (HM 48-1), CD48 PE (HM 48-1), CD117 eFluor780 (2B8), CD127 PE-Cy7 (A7R34), CD135 PE (A2F10), Sca-1 PE-Cy7 (D7), Sca-1 PerCP-Cy5.5 (D7), Tie2 biotin (bio) (TEK4), Tie2 PE (TEK4) (eBioscience), CD3e bio (500A2), CD4 bio (GK1.5), CD4 PE (H129.19), CD8 bio (53-6.7), CD8 PE (53-6.7), CD11b APC (M1/70), CD19 APC (ID3), CD19 bio (ID3), CD19 PE (ID3), CD45 bio (30-F11), Gr-1 APC (RB6-8C5), Gr-1 bio (RB6-8C5), Gr-1 PE (RB6-8C5), IgM bio (R6-60.2), IgM PE (R6-60.2), Streptavidin PE-Cy7, Sca-1 PE (E13 161.7), Ter119 APC (Ter119), Ter119 bio (Ter119), Ter119 PE (Ter119) (BD Pharmingen), CD4 APC (RM4.5), CD19 QDot605 (6D9), streptavidin APC, streptavidin QDot605, Sca-1 PE-Cy5.5 (D7) (Invitrogen/Molecular Probes), CD11b bio (M1/70.15) (Caltag), CD135 bio (A2F19), CD150 PE-Cy7 (TC15-12F12.2), CD229 PE (Ly9AB3) (Biolegend). The lineage cocktail (Lin) was composed of CD3e, CD4, CD8, CD11b, CD19, Gr-1 and Ter119. To enrich for Lin-negative progenitor populations, bone marrow cells were stained with lineage markers followed by depletion with Dynabeads (Life technologies) according to the manufacturer's instruction. Dead cells were excluded by staining with Sytox Blue (Invitrogen). Cells were analysed on a FACSFortessa, or sorted by FACSAriaIII (all

Becton & Dickinson), and data were analysed by BD FACSDiva software. Bone marrow cell populations were defined as follows: HSCs (Lin⁻ Kit⁺ Sca-1⁺ CD150⁺ CD48⁻), ST-HSCs (Lin⁻ Kit⁺ Sca-1⁺ CD150⁺ CD48⁻), ST-HSCs CD229⁺, ST-HSCs CD229⁺, MPPs (Lin⁻ Kit⁺ Sca-1⁺ CD150⁺ CD48⁺), CLPs (Lin⁻ IL7R⁺ Flk2⁺ Kit^{low} Sca-1^{low}), pro B cells (CD19⁺ IgM⁻), B cells (CD19⁺ IgM⁺), CMPs (Lin⁻ Kit⁺ Sca-1⁻ CD34⁺ CD16/32^{low}), GMPs (Lin⁻ Kit⁺ Sca-1⁻ CD34⁺ CD16/32⁺) and MEPs (Lin⁻ Kit⁺ Sca-1⁻ CD34⁺ CD16/32⁻). Populations in the spleen were defined as follows: T cells (CD3⁺ CD19⁻ CD11b⁻ Gr-1⁻), B cells (CD19⁺ CD3⁺ CD11b⁻ Gr-1⁻), granulocytes (Gr-1⁺ CD11b⁺ CD3⁻ CD19⁻) and macrophages (CD11b⁺ Gr-1⁻ CD3⁻ CD19⁻). The populations in the thymus were defined as follows: double-negative thymocytes (CD4⁻ CD8⁻) (DN), double-positive thymocytes (CD4⁺ CD8⁺) (DP), CD4⁺ thymocytes (CD4⁺ CD8⁻) (CD4) or CD8⁺ thymocytes (CD4⁻ CD8⁺) (CD8). Erythroid progenitors (Lin⁻ Kit⁺ CD45^{low}) in fetal liver were defined as described previously^{21,32}.

Mice. Mice carrying constitutively active reporter alleles (*panYFP* or *panRFP*) were generated by crossing *Rosa*^{YFP} or *Rosa*^{RFP} reporter mice^{31,33} to germline Cre-deleter mice³⁴. Offspring constitutively expressing YFP or RFP in all tissues were backcrossed towards the C57BL/6 background, and used in competitive transplantation experiments (Extended Data Fig. 2h, i) and in adoptive transfers of myeloid and lymphoid progenitors (Extended Data Fig. 4). *Rag2*^{-/-} γ c^{-/-} Kit^{W/Wv} mice¹⁹ were used for adoptive transfer experiments without previous irradiation (Extended Data Figs 1f, 2h, i and 4). For transfer experiments with irradiation, congenic B6.SJL-Ptprca Pep3b/BoyJ (H-2^b; CD45.1⁺) mice were used as recipients (Fig. 5). All animal procedures were approved by the Regierungspräsidium Karlsruhe, and performed in accordance with the Institutional Guidelines. Both male and female mice were used at ages ranging from embryonic E12.5 to around 120 weeks (2 years). No mice were excluded from the analysis. No randomization and no blinding were used.

PCR genotyping. Tissues were lysed in lysis buffer (DirectPCR-Lysis Reagent Tail, Peqlab) according to the manufacturer's instruction. Mice were genotyped by PCR for 2 min at 94 °C (20 s at 94 °C, 30 s at 51 °C, 1 min at 72 °C) 35 times; 10 min at 72 °C using a common 5' oligonucleotide annealing upstream of the rabbit β -hemoglobin gene (5'-CATCGCATACCATACATAGGTGGAGG-3') and a 3' oligonucleotide annealing to the rabbit β -hemoglobin gene (5'-AATCAAGGGTCC CCAACTCAC-3'), yielding a 526-bp DNA fragment indicating the *Tie2*^{MCM} allele, and a 3' oligonucleotide (5'-GAGGCAGCATCTGTCTACAAGAGATGG-3'), yielding a 745-bp DNA fragment indicating the *Tie2*⁺ allele.

Single-cell transplantation. YFP⁺ HSCs (LSK CD150⁺ CD48⁻) were isolated from tamoxifen-treated *Tie2*^{MCM/+} *Rosa*^{YFP} mice by electronic single-cell deposition into individual wells of a U-bottom 96-well plate containing 100 μ l sterile 5% FCS. Before injection, 100 μ l sterile PBS was added to each well, and single cells were injected intravenously into individual *Rag2*^{-/-} γ c^{-/-} Kit^{W/Wv} recipients. Peripheral blood samples were collected 4–8 weeks after transfer from the submandibular vein into EDTA-containing microtubes (Sarstedt) to screen for progeny of donor cells (detected by YFP expression). Organs of recipient mice were analysed at least 16 weeks after transplantation as shown in Extended Data Fig. 1f. For secondary transplantations, YFP⁺ Lin⁻ Kit⁺ bone marrow cells from the primary recipients were purified by cell sorting and 3×10^4 cells were intravenously injected into individual *Rag2*^{-/-} γ c^{-/-} Kit^{W/Wv} recipients. At least 16 weeks after transfer organs were analysed by flow cytometry.

Competitive transplantation. Equal cell numbers (500–1,000 of each) of HSCs (LSK CD150⁺ CD48⁻) from *Tie2*^{MCM/+} mice (either with *panRFP* reporter allele or without), and wild-type competitor HSCs (LSK CD150⁺ CD48⁻) from *Tie2*^{+/+} mice (*panYFP*) were co-injected intravenously into *Rag2*^{-/-} γ c^{-/-} Kit^{W/Wv} recipient mice. Peripheral blood samples were collected 4–8 weeks after transfer to screen for progeny of donor cells (not shown), and after at least 16 weeks the organs indicated in Extended Data Fig. 2 were analysed for the contributions of progeny from *Tie2*^{MCM/+} or *Tie2*^{+/+} HSC. To assess the life span of myeloid progenitor and mature granulocytes, cell-sorter-purified CMPs and GMPs (mixed as one population) from *panRFP* mice were intravenously injected together with CLPs from *panYFP* mice into *Rag2*^{-/-} γ c^{-/-} Kit^{W/Wv} recipient mice (5×10^4 CMPs plus GMPs and 0.5×10^4 CLPs per mouse). Peripheral blood samples were collected 7, 14, 21 and 32 days after transplantation from the submandibular vein into EDTA-containing microtubes (Sarstedt) to screen by flow cytometry for donor-derived progeny cells.

Transplantation into lethally irradiated mice. Lin-negative Kit⁺ cells from tamoxifen-treated *Tie2*^{MCM/+} *Rosa*^{YFP} mice (CD45.2) were sorted and 0.5×10^5 – 2.0×10^5 cells injected intravenously into lethally irradiated (1,100 cGy; split dose with 4 h time gap between each dose; Cesium 137 GammaCell40 Irradiator, Besttheratronics) congenic B6.SJL-Ptprca Pep3b/BoyJ mice. Recipient mice were maintained on antibiotic water (1.17 g l⁻¹ neomycin sulphate) for 14 days. Bone marrow of recipient mice was analysed after at least 16 weeks after transplantation.

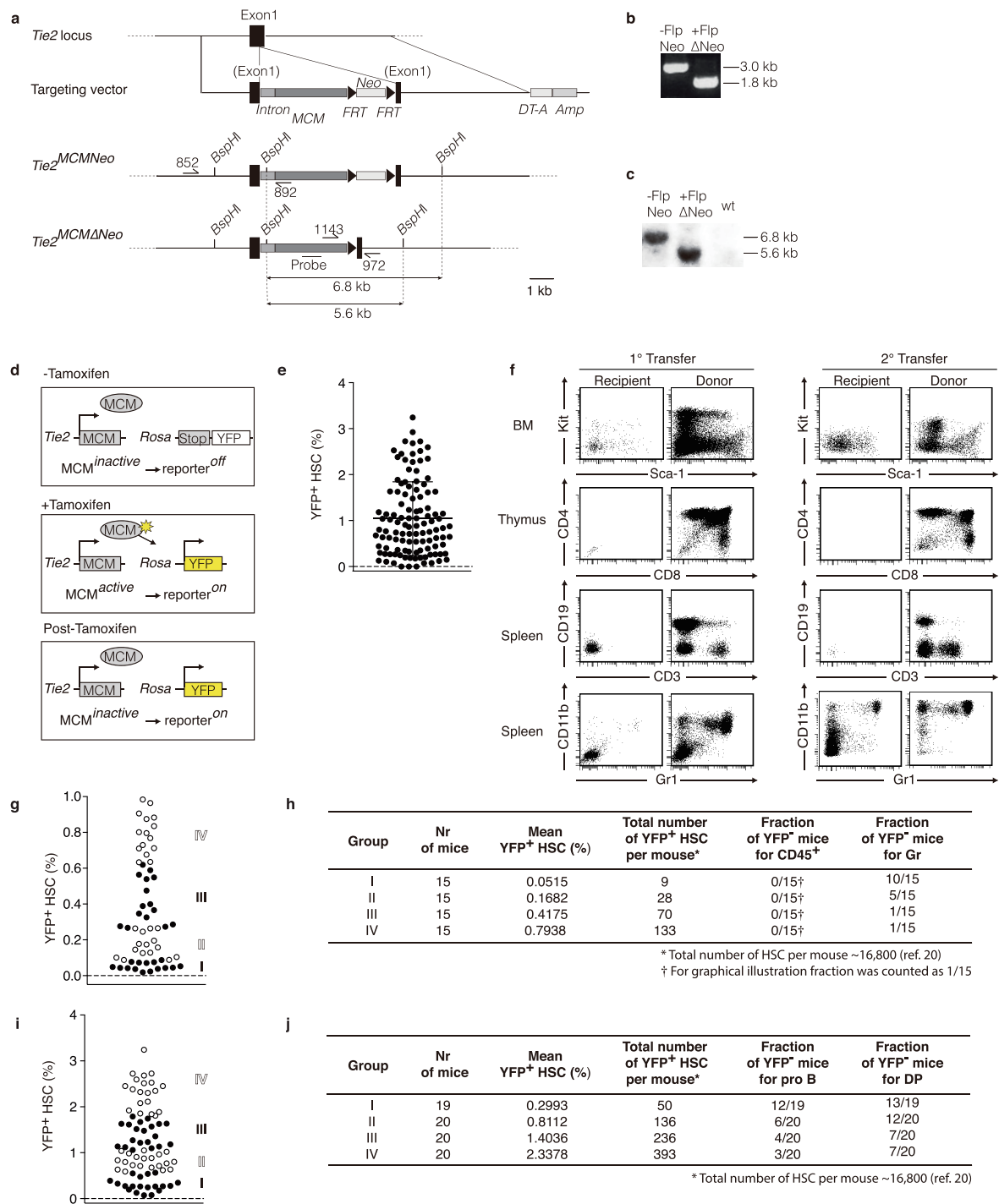
Proliferation assay. *Tie2*^{MCM/+} and *Tie2*^{+/+} littermate mice were injected intraperitoneally with 1 mg EdU (Invitrogen) in PBS. After 24 h, the bone marrow was

collected, cells were stained with the appropriate antibodies before sorting cell-surface receptor Tie2-positive (LSK Tie2⁺CD150⁺CD48⁻) and cell-surface receptor Tie2-negative (LSK Tie2⁻CD150⁺CD48⁻) HSCs, each from Tie2^{MCM/+} and Tie2^{+/+} littermates. Cells were collected in individual tubes containing 50% FCS, fixed and permeabilized, and the Click-it reaction was performed using the Click-iT EdU Flow Cytometry Assay Kit (Invitrogen) according to the manufacturer's protocol.

5-FU treatment. 5-FU (Sigma F6627) was dissolved in sterile PBS, and tamoxifen-treated Tie2^{MCM/+}Rosa^{YFP} mice were intravenously injected with a single dose of 250 mg kg⁻¹ or PBS. Mice were maintained on antibiotic (1.17 g neomycin sulphate per litre of drinking water) for 14 days. Peripheral blood was collected 4, 7, 10 and 12 days after 5-FU from the submandibular vein into EDTA-containing microtubes (Sarstedt). Absolute numbers of leukocytes per microlitre blood were determined by flow cytometry using anti CD45 PE-Cy7 (30-F11; eBioscience) antibody and APC-conjugated CaliBRITE Beads (BD Biosciences) as standard. On day 12 or 18 after 5-FU treatment bone marrow cells were isolated and analysed by flow cytometry. With two-tailed *t*-test, effect size $d = 1$, $\alpha = 0.05$ and power 0.8, group sizes should be 17 per sample group. We chose $n = 15$ (control) and $n = 18$ (5-FU treatment).

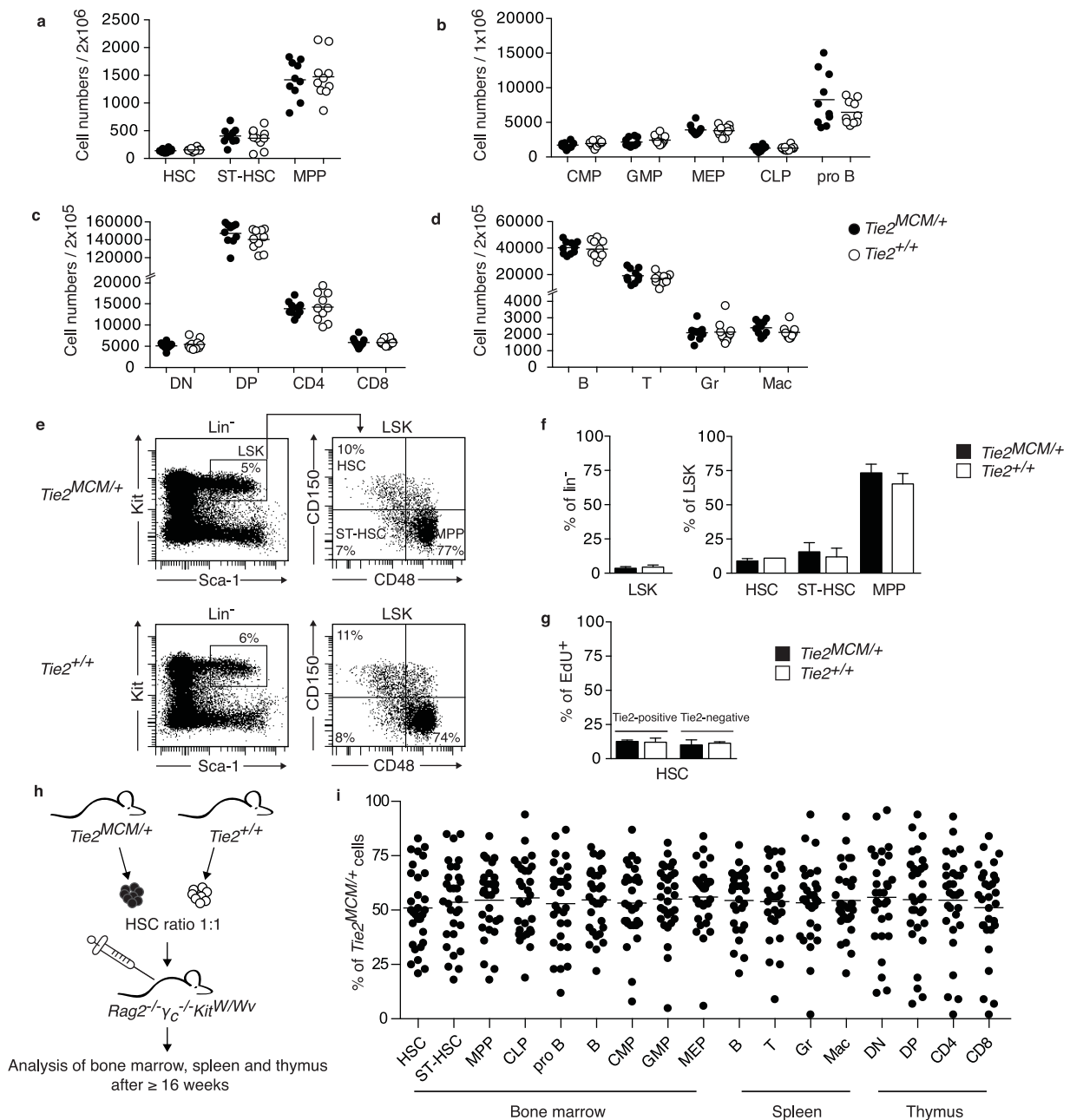
Mathematical methods. Mathematical modelling and parameter inference from the experimental data are described in the Supplementary Information. The corresponding Matlab codes are available on request. No statistical methods were used to predetermine sample size, except for 5-FU treatment.

25. Kouskoff, V., Fehling, H. J., Lemeur, M., Benoist, C. & Mathis, D. A vector driving the expression of foreign cDNAs in the MHC class II-positive cells of transgenic mice. *J. Immunol. Methods* **166**, 287–291 (1993).
26. Shimshek, D. R. *et al.* Codon-improved Cre recombinase (iCre) expression in the mouse. *Genesis* **32**, 19–26 (2002).
27. Verrou, C., Zhang, Y., Zurn, C., Schamel, W. W. & Reth, M. Comparison of the tamoxifen regulated chimeric Cre recombinases MerCreMer and CreMer. *Biol. Chem.* **380**, 1435–1438 (1999).
28. Casanova, E. *et al.* ER-based double iCre fusion protein allows partial recombination in forebrain. *Genesis* **34**, 208–214 (2002).
29. Dumont, D. J. *et al.* Dominant-negative and targeted null mutations in the endothelial receptor tyrosine kinase, *tek*, reveal a critical role in vasculogenesis of the embryo. *Genes Dev.* **8**, 1897–1909 (1994).
30. Sato, T. N. *et al.* Distinct roles of the receptor tyrosine kinases Tie-1 and Tie-2 in blood vessel formation. *Nature* **376**, 70–74 (1995).
31. Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4 (2001).
32. Bertrand, J. Y. *et al.* Three pathways to mature macrophages in the early mouse yolk sac. *Blood* **106**, 3004–3011 (2005).
33. Luche, H., Weber, O., Nageswara Rao, T., Blum, C. & Fehling, H. J. Faithful activation of an extra-bright red fluorescent protein in “knock-in” Cre-reporter mice ideally suited for lineage tracing studies. *Eur. J. Immunol.* **37**, 43–53 (2007).
34. Schwenk, F., Baron, U. & Rajewsky, K. A cre-transgenic mouse strain for the ubiquitous deletion of *loxP*-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23**, 5080–5081 (1995).



Extended Data Figure 1 | Generation of the *Tie2^{MCM}* allele, labelling of HSCs in tamoxifen-treated *Tie2^{MCM/+} Rosa^{YFP}* mice, and groups for limiting dilution analysis. **a, The endogenous *Tie2* locus, the gene targeting vector and the targeted allele with (*Tie2^{MCMNeo}*) and without (*Tie2^{MCMΔNeo}*) neomycin are depicted. Oligonucleotides and DNA probe for genotyping, and restriction sites used for Southern blotting are indicated (not drawn to scale). **b**, PCR verification of the targeted *Tie2^{MCM}* allele in embryonic stem cells before and after neomycin deletion. **c**, Southern blot analysis of *Tie2^{MCMNeo}* (–Flp) and *Tie2^{MCMΔNeo}* (+Flp), and wild-type embryonic stem-cell clones. **d**, Principle of inducible fate mapping. In the absence of tamoxifen MCM is inactive (reporter^{off}). Tamoxifen treatment activates the reporter (reporter^{on}). After tamoxifen treatment, labelled cells and their progeny remain marked (reporter^{on}). **e**, Summary of HSC labelling frequencies of tamoxifen-treated *Tie2^{MCM/+} Rosa^{YFP}* mice ($n = 112$; 5 times on 5 consecutive days) analysed between 1 and 34 weeks after label induction. These data are the basis for the kinetic analysis (Fig. 2a–f) and for the mathematical modelling (Fig. 3). Each**

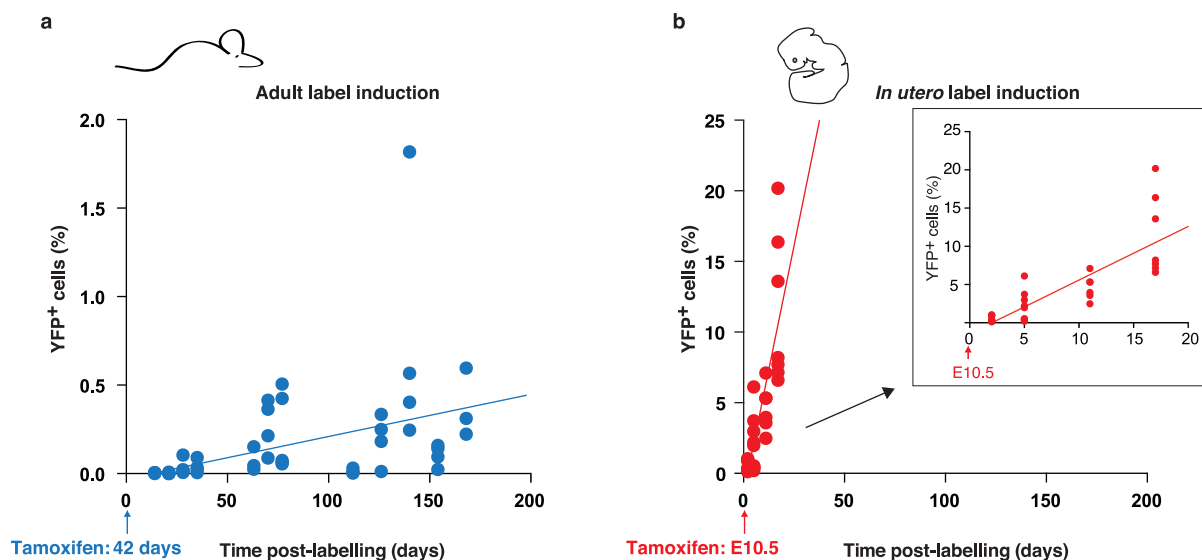
dot represents an individual mouse. Bar indicates mean (1.041 ± 0.8013 s.d.). **f**, A single YFP⁺ LSK CD150⁺ CD48[–] HSC from a tamoxifen-treated *Tie2^{MCM/+} Rosa^{YFP}* mouse was transplanted into a *Rag2^{–/–} γc^{–/–} Kit^{W/W^v}* recipient mouse (1° transfer). Donor cells were identified by YFP expression, and analysed 16 weeks after transplantation in bone marrow (BM), thymus and spleen using the markers shown. YFP⁺ Kit⁺ donor bone marrow cells were re-transplanted into a secondary *Rag2^{–/–} γc^{–/–} Kit^{W/W^v}* recipient (2° transfer), and analysed as described for the primary transfer. **g–j**, HSC labelling frequencies in tamoxifen-treated *Tie2^{MCM/+} Rosa^{YFP}* mice analysed 6 weeks onwards after label induction were used for limiting dilution analysis of CD45⁺ output, granulocytes ($n = 60$) (**g, h**), pro B cells and double-positive thymocytes ($n = 79$) (**i, j**). Each dot represents an individual mouse. Mice grouped together are highlighted in black or white (groups I–IV). Mathematical calculations are shown in the tables (**h, j**). In **g**, data shown represent the aggregate of labelling frequencies below 1% shown in **e**, plus data obtained in mice receiving only a single tamoxifen injection.



Extended Data Figure 2 | Characterization of *Tie2*^{MCM/+} mice.

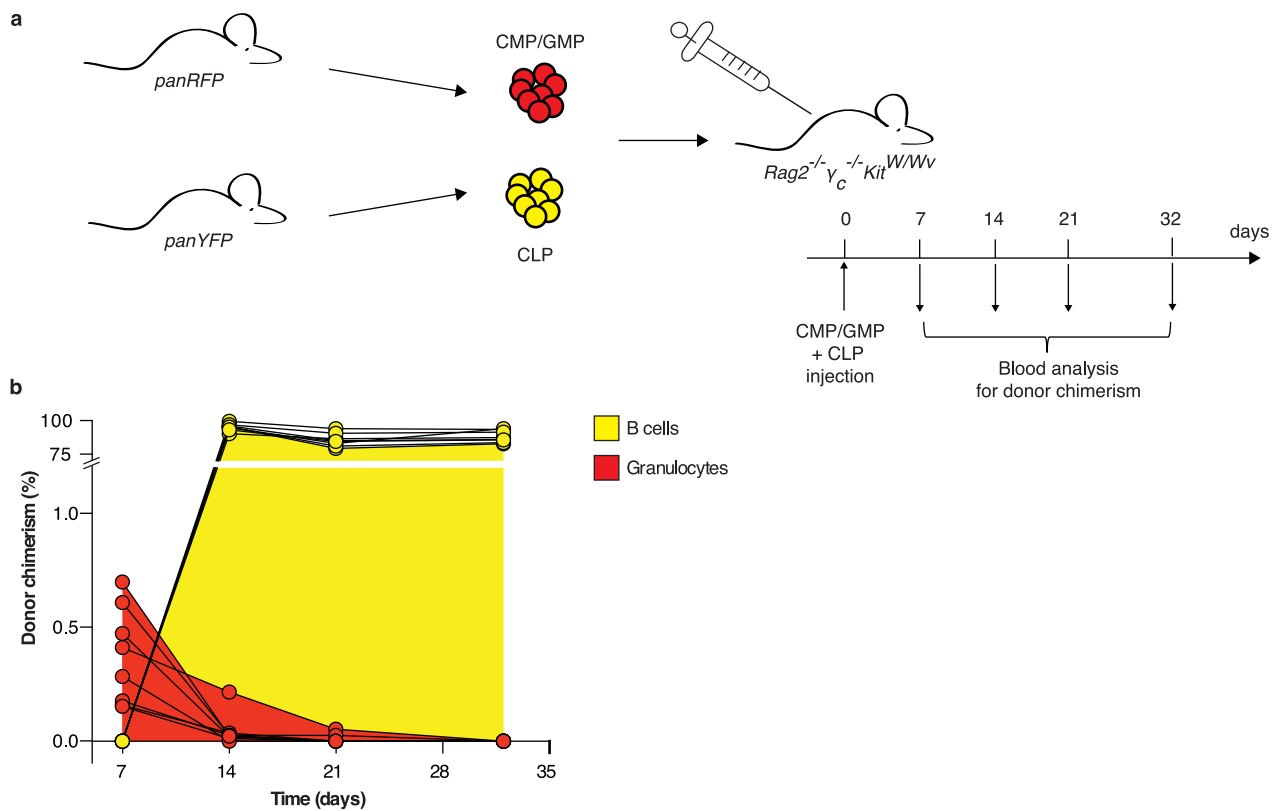
a–d, Numbers of haematopoietic cell subsets isolated from bone marrow (**a**, **b**), thymus (**c**) and spleen (**d**) of *Tie2*^{MCM/+} ($n = 10$; black) and *Tie2*^{+/+} ($n = 10$; white) littermates were determined by flow cytometry. Each dot represents an individual mouse. Bars indicate mean. **e**, **f**, Flow cytometric analysis of a representative *Tie2*^{MCM/+} (top) and *Tie2*^{+/+} (bottom) mouse gated on Lin⁻ cells, and analysed for expression of Kit versus Sca-1. Further analysis of Lin⁻Kit⁺Sca-1⁺ (LSK) cells for CD150 and CD48 revealed comparable marker distributions (**f**). Percentages of LSK cells among the Lin⁻ fraction (left), and of HSCs, ST-HSCs and MPPs among the LSK fraction (right) in the bone marrow of three independent *Tie2*^{MCM/+} (black) and

Tie2^{+/+} (white) mice. Data are mean \pm s.d. **g**, Proliferation rates in surface receptor Tie2-positive and Tie2-negative HSCs in the bone marrow of *Tie2*^{MCM/+} (black) and *Tie2*^{+/+} (white) mice 24 h after EdU administration. Data represent mean \pm s.d. from two independent experiments of FACS-sorted populations from *Tie2*^{MCM/+} ($n_{\text{Exp1}} = 5$; $n_{\text{Exp2}} = 3$) and *Tie2*^{+/+} ($n_{\text{Exp1}} = 3$; $n_{\text{Exp2}} = 2$) mice. **h**, **i**, *Rag2*^{-/-} γ _c^{-/-}Kit^{W/W^v} recipients ($n = 30$; for analysis of B and T cells in the spleen $n = 28$) were injected with equivalent numbers of *Tie2*^{MCM/+} and *Tie2*^{+/+} HSCs (500–1,000 of each), and analysed after at least 16 weeks. The percentages of *Tie2*^{MCM/+} HSC-derived haematopoietic cells in bone marrow, spleen and thymus are shown. Each dot represents an individual mouse. Bars indicate mean.



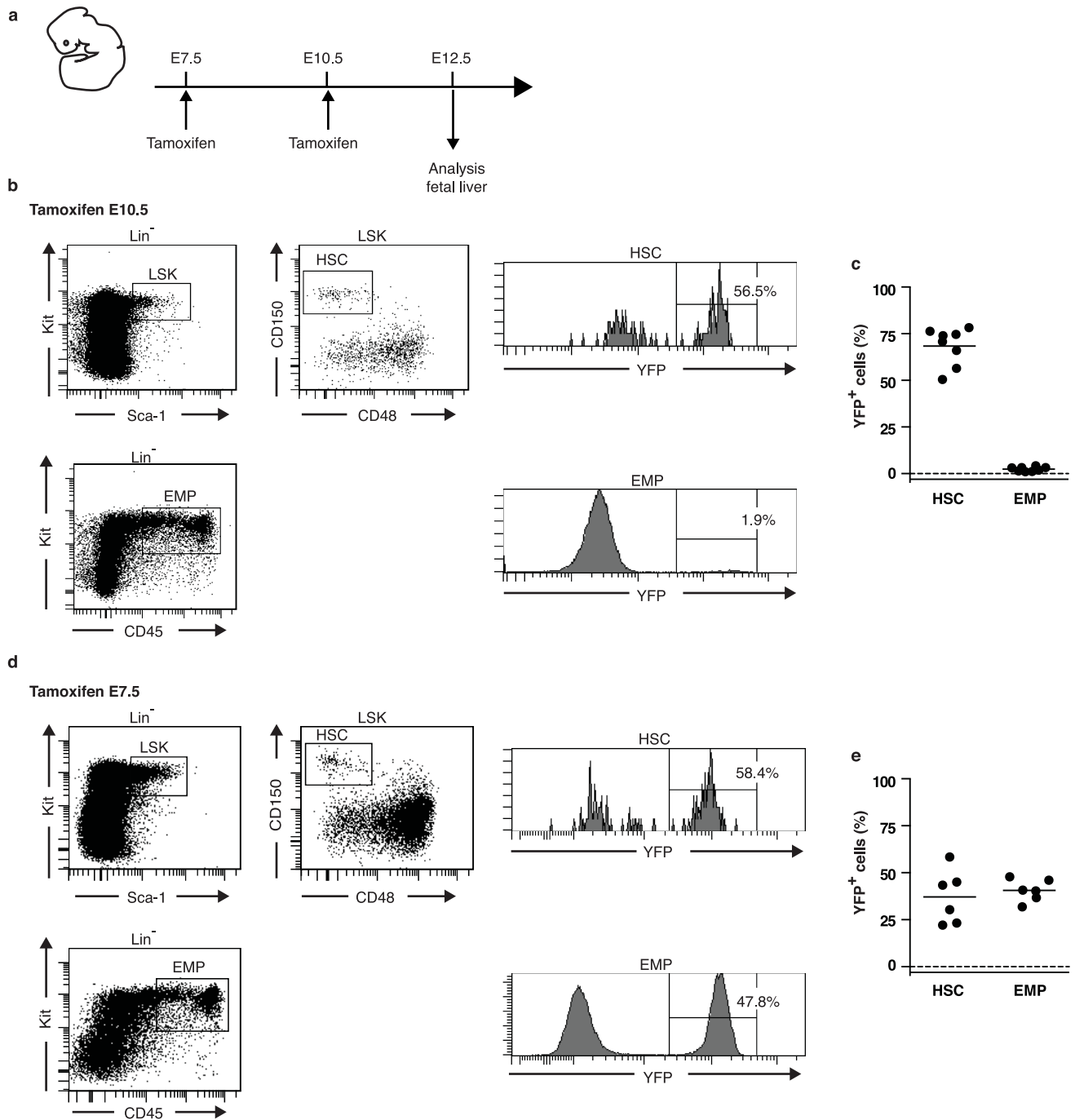
Extended Data Figure 3 | Kinetics of YFP label emergence after label induction in total bone marrow cells in adult mice, and in fetal liver and bone marrow cells in fetal and early postnatal mice. a, Percentages of YFP⁺ cells among total non-lineage-depleted bone marrow cells of tamoxifen-treated *Tie2^{MCM/+}Rosa^{YFP}* mice ($n = 47$). Time point 0 corresponds to the time of

tamoxifen treatment of adult mice (all of which were at least 6 weeks). **b,** *Tie2^{MCM/+}Rosa^{YFP}* mice ($n = 32$) were treated with tamoxifen on E10.5 (time point 0). Subsequently, percentages of YFP⁺ cells were determined among total fetal liver on E12.5 and E15.5, and in bone marrow at birth and 1 week of age. Each dot represents an individual mouse.



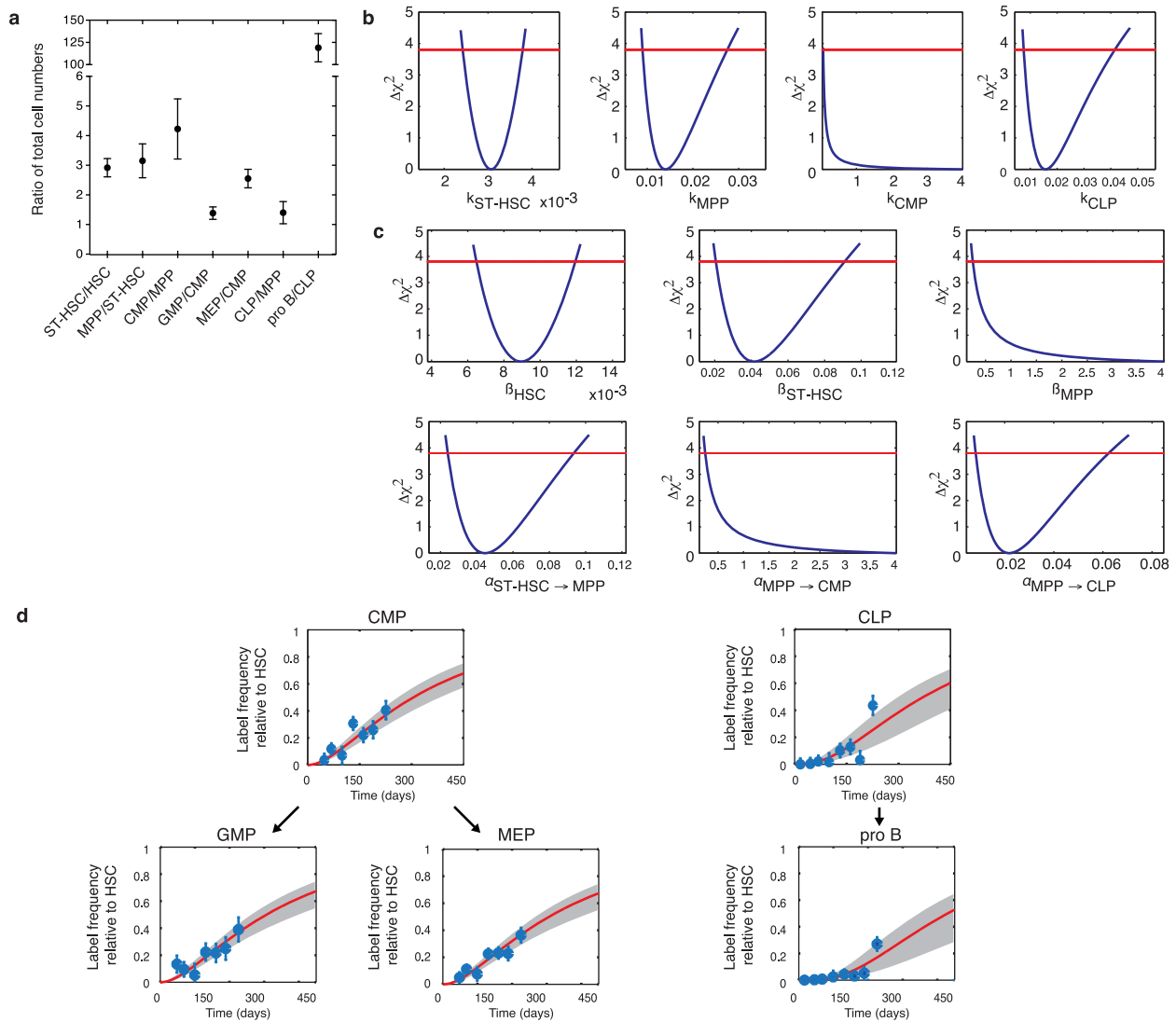
Extended Data Figure 4 | Donor-derived myeloid cells disappear within 6 weeks after adoptive transfer. **a**, *Rag2^{-/-} γ_c^{-/-} Kit^{W/Wv}* recipient mice ($n = 8$) received CMPs and GMPs (together 5×10^4 per mouse) from *panRFP* mice together with CLPs (0.5×10^4 per mouse) from *panYFP* mice. Peripheral

blood samples were measured 7, 14, 21 and 32 days after transplantation. **b**, Analysis of donor chimaerism of CMP/GMP-derived ($\text{Gr-1}^+ \text{CD11b}^+$ granulocytes, red) and CLP-derived (CD19^+ B cells, yellow) progeny. Each line represents an individual mouse.



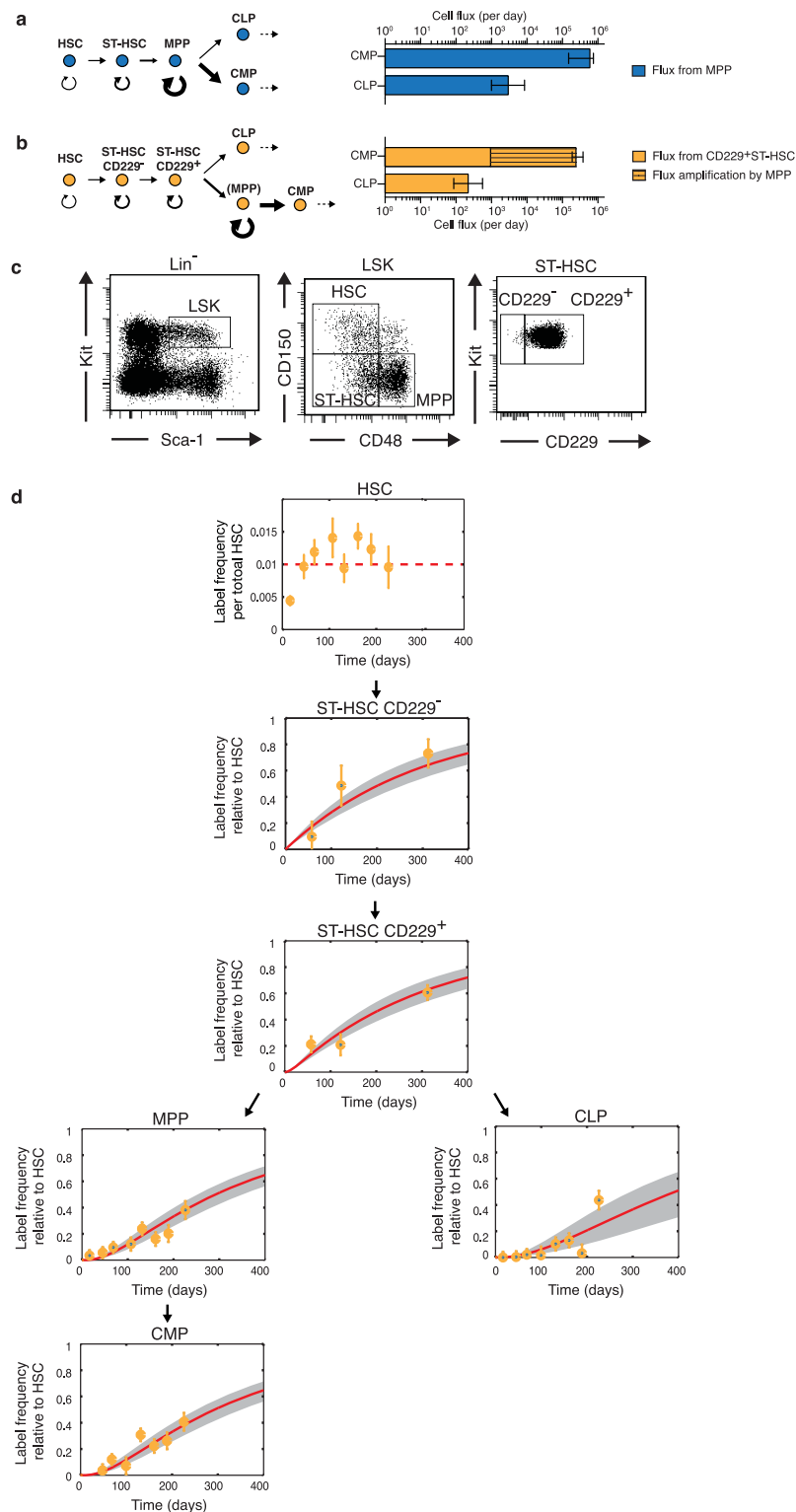
Extended Data Figure 5 | Labelling of HSCs, but not erythromyeloid progenitors, in *Tie2^{MCM/+}Rosa^{YFP}* embryos treated on E10.5 *in utero* by tamoxifen. **a**, *Tie2^{MCM/+}Rosa^{YFP}* embryos were treated with tamoxifen on E7.5 or E10.5, and analysed at E12.5. **b–e**, Fetal liver cells from representative *Tie2^{MCM/+}Rosa^{YFP}* embryos labelled on E10.5 (**b**) or E7.5 (**d**) were analysed by flow cytometry for YFP labelling in HSCs (LSK CD150⁺CD48⁺) and

erythromyeloid progenitors (EMP) (Lin[−]Kit⁺CD45^{+/low}) (see Methods for references). HSCs were marked in mice labelled at E7.5 and E10.5, but erythromyeloid progenitors were not marked in mice labelled on E10.5. **c, d**, Summary of experiments depicted in **b** ($n = 8$) (**c**) and **d** ($n = 6$) (**e**). Each dot represents an individual embryo. Bars indicate mean.



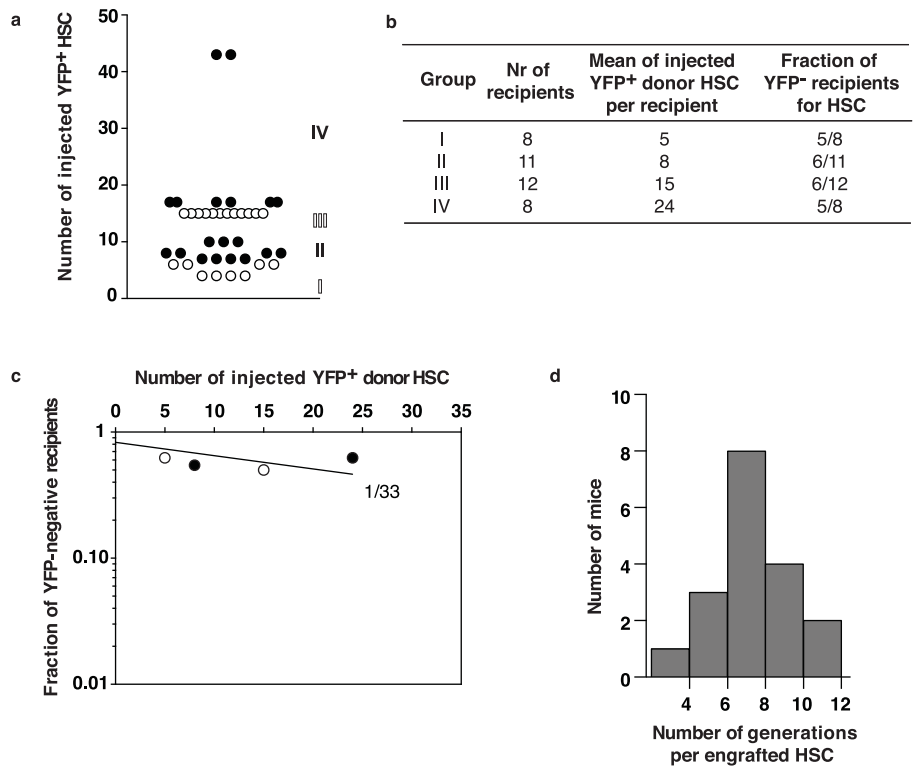
Extended Data Figure 6 | Details of model analysis. **a**, Ratio of compartment sizes for stem and progenitor cell compartments. Compartment sizes were determined by cell counting of cell suspensions from bone marrow, spleen and thymus in a Neubauer chamber combined with phenotypic definition of each population by flow cytometry. Mean and s.e.m. for $n = 10$ mice are shown. **b**, Confidence bounds for the model parameters. Profile likelihoods (blue lines) for the inverse residence times, $k = 1/\tau$, of cells in the ST-HSC, MPP, CMP and CLP compartments (units: per day). The profile likelihoods have been computed as described in Supplementary Methods, with the experimental data from Fig. 3c. The red lines indicate the 95% confidence level. Except for the CMPs, for which the data are consistent with a broad range of k (that is, the CMP residence time is low), the residence times are accurately determined by the label propagation data alone. In particular, the very low k for the ST-HSCs shows that this compartment operates near self-renewal. **c**, Profile likelihoods (blue lines) and 95% confidence level (red lines) for the net proliferation rates (β) and differentiation rates (α), computed with the data of Fig. 3c and Extended Data Fig. 4 (unit: per day). The differentiation rate from HSCs ($\alpha_{\text{HSC} \rightarrow \text{ST-HSC}}$) has the same profile likelihood as the net proliferation of HSC

(β_{HSC}) and is therefore not shown. $\alpha_{\text{HSC} \rightarrow \text{ST-HSC}}$, $\alpha_{\text{ST-HSC} \rightarrow \text{MPP}}$, $\alpha_{\text{MPP} \rightarrow \text{CLP}}$, β_{HSC} and $\beta_{\text{ST-HSC}}$ are accurately determined by the data. Moreover, both $\alpha_{\text{MPP} \rightarrow \text{CMP}}$ and β_{MPP} have a lower bound that is one and two orders of magnitude larger than the respective parameters for ST-HSCs and HSCs, showing that MPPs have significantly higher proliferation and differentiation activities than the preceding compartments. Note that (β_{MPP}) and $\alpha_{\text{MPP} \rightarrow \text{CMP}}$ are strongly correlated (not shown). **d**, Label progression data and fit of the mathematical model for further downstream myeloid precursors (CMPs to GMPs or MEPs) and lymphoid precursors (CLPs to pro B cells) in the bone marrow. Data were measured up to 238 days after label induction (blue points, with s.e.m., as in Fig. 2f) have been used for the model fit (red lines, best fit and grey shades, 95% confidence bands). The parameter values are as in Supplementary Table 1 and, in addition, $\alpha_{\text{CMP} \rightarrow \text{GMP}} = 2$ (0.04, 4) day^{-1} , $\alpha_{\text{CMP} \rightarrow \text{MEP}} = 3$ (0.1, 4) day^{-1} , $\alpha_{\text{CLP} \rightarrow \text{pro B}} = 2$ (0.8, 4) day^{-1} , $\beta_{\text{CMP}} = 4$ (−1, 4) day^{-1} , $\beta_{\text{CLP}} = 3$ (0.4, 4) day^{-1} , $\tau_{\text{GMP}} = 0.12$ (0.12, 33) days, $\tau_{\text{MEP}} = 0.13$ (0.13, 22) days, $\tau_{\text{proB}} = 54$ (6, 141) days (in brackets: 95% confidence bounds). For further details see Supplementary Information.



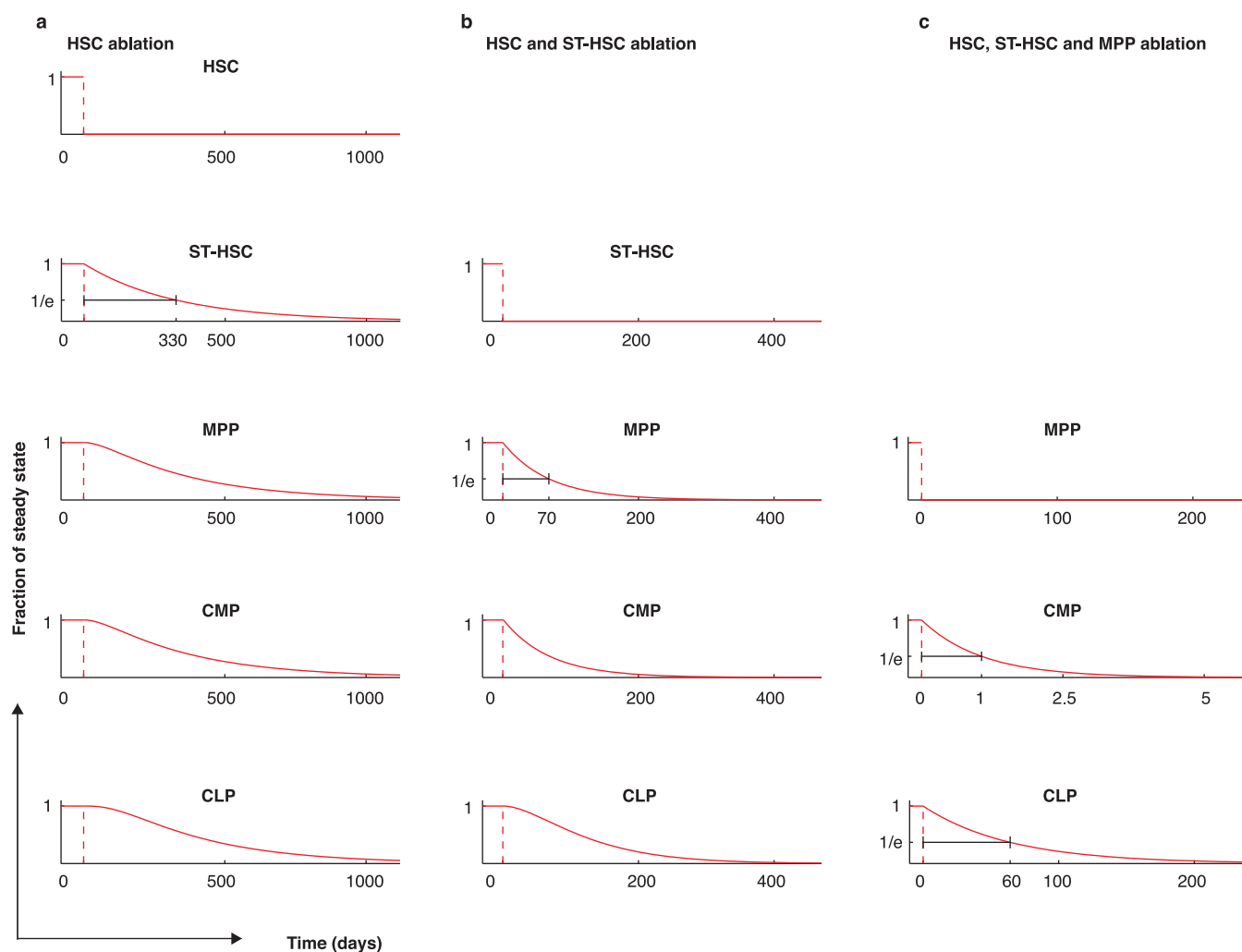
Extended Data Figure 7 | Consequences of lymphoid-myeloid branch points at distinct stem or progenitor stages. **a–c**, Lymphoid-myeloid branch points were considered at the MPP stage (**a**) or at the earlier CD229⁺ ST-HSC subset (**b**). Each population is phenotypically defined as shown in **c**. MPPs have also been termed HPC-1, and the CD229⁺ ST-HSC subset has also been termed MPP2/3 (ref. 18). Cell flux rates per day are shown from the branch points to CMPs or CLPs (right panels in **a** and **b**, with 95% confidence intervals). In both scenarios, the production of CMPs is several-hundred-fold larger than that of CLPs. Assuming the branch point at the CD229⁺ ST-HSC stage, biased myeloid differentiation is still evident (~5-fold), but the large uneven production is mainly achieved by flux amplification downstream from the bifurcation. **d**, Label progression data and fit of the mathematical model.

Data measured up to 311 days after label induction in the ST-HSC CD229⁻ and ST-HSC CD229⁺ compartments (orange points, with s.e.m.; groups of mice: 57 days $n = 7$; 122 days $n = 4$; 311 days $n = 8$) and in CLP, MPP and CMP compartments data measured up to 238 days (orange points, with s.e.m.; groups of adult mice as in Fig. 2f) have been used for the model fit (red lines, best fit; grey shades, 95% confidence bands). The resulting parameters are $\alpha_{\text{HSC} \rightarrow \text{CD229}^- \text{ST-HSC}} = 0.0001$ (0.00001–0.00016) day^{-1} , $\alpha_{\text{CD229}^- \text{ST-HSC} \rightarrow \text{CD229}^+ \text{ST-HSC}} = 4$ (2–4) day^{-1} , $\alpha_{\text{CD229}^+ \text{ST-HSC} \rightarrow \text{MPP}} = 0.03$ (0.02–0.06) day^{-1} , $\alpha_{\text{CD229}^+ \text{ST-HSC} \rightarrow \text{CLP}} = 0.007$ (0.004–0.012) day^{-1} , $\beta_{\text{CD229}^- \text{ST-HSC}} = 4$ (2–4) day^{-1} , $\beta_{\text{CD229}^+ \text{ST-HSC}} = 0.0$ (–0.06–0.0016) day^{-1} and $\beta_{\text{MPP}} = 4$ (0.3–4) day^{-1} (in brackets: 95% confidence intervals). For further details see Supplementary Methods.



Extended Data Figure 8 | Limiting dilution analysis of *in-situ*-labelled transplanted HSCs. **a**, *Tie2^{MCM/+}Rosa^{YFP}* mice were injected with tamoxifen (donor bone marrow; see Fig. 5a). Numbers of YFP⁺ HSCs in donor bone marrow transplanted into 38 recipient mice are displayed. Each dot is an individual recipient. Mice grouped together are highlighted in black or white (groups I–IV). **b**, Data underlying limiting dilution analysis. **c**, Numbers of injected YFP⁺ donor HSCs are plotted against the fraction of YFP-negative

recipients on a logarithmic scale. **d**, Histogram showing the extent of self-renewing proliferation for engrafted YFP⁺ HSCs in mice labelled ‘within sampling error’ and ‘overrepresented’ in Fig. 5d ($n = 18$). Proliferation is shown as the number of generations per engrafted HSC needed to achieve the measured frequencies of YFP⁺ HSCs in the bone marrow 4 months after transplantation. For further details see Supplementary Discussion.



Extended Data Figure 9 | Simulated effects of ablation of HSCs, ST-HSCs or MPPs. In the mathematical model used to fit the label progression data (Fig. 3c), numbers of HSCs (a), HSCs and ST-HSCs (b), and HSCs, ST-HSCs and MPPs (c) were reduced from its steady state value to zero at $t = 0$, and subsequently held there. The predicted responses in the downstream

compartments are shown. The horizontal black lines at $1/e = 37\%$ indicate the residence times of the compartments that control the time scale of the response. Note that these simulations assume that there are no homeostatic mechanisms present within the compartments that could maintain their size independent of input from upstream compartments.

Extended Data Table 1 | Transplantations of *in-situ*-labelled HSC

Fraction	Number of injected donor cells	Number of 1° recipients	1° recipients		Number of 2° recipients	2° recipients	
			Long-term	Transient		Long-term	Transient
LSK CD150+ CD48- YFP+	1	23	1(4%)	4(17%)	3	1(33%)	2(66%)
LSK CD150+ CD48- YFP+	5	1	1(100%)	-	n.d.		
LSK CD150+ CD48- YFP+	8	1	1(100%)	-	n.d.		
LSK YFP+	10	3	1(33%)	2(66%)	3	1(33%)	2(66%)
LSK YFP+	15	3	1(33%)	2(66%)	3	1(33%)	2(66%)
LSK YFP+	18	4	1(33%)	-	3	1(33%)	2(66%)
LSK YFP+	20	3	1(33%)	2(66%)	5	2(40%)	3(60%)
LSK YFP+	40	3	3(100%)	-	9	2(22%)	7(88%)

n.d.: not done

YFP-marked stem cells from tamoxifen-treated *Tie2^{MCM/+} Rosa^{YFP}* mice were injected into primary and secondary *Rag2^{-/-} γc^{-/-} Kit^{W/W}* recipient mice. Phenotypes and cell numbers of injected cells, numbers of primary (1°) and secondary (2°) recipients, and reconstitution results are given. Percentages in brackets reflect the proportions of reconstituted mice.

Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors

Elisa Gomez Perdiguero^{1*}, Kay Klapproth^{2*}, Christian Schulz¹, Katrin Busch², Emanuele Azzoni³, Lucile Crozet¹, Hannah Garner¹, Celine Trouillet¹, Marella F. de Bruijn³, Frederic Geissmann^{1§} & Hans-Reimer Rodewald^{2§}

Most haematopoietic cells renew from adult haematopoietic stem cells (HSCs)^{1–3}, however, macrophages in adult tissues can self-maintain independently of HSCs^{4–7}. Progenitors with macrophage potential *in vitro* have been described in the yolk sac before emergence of HSCs^{8–13}, and fetal macrophages^{13–15} can develop independently of *Myb*⁴, a transcription factor required for HSC¹⁶, and can persist in adult tissues^{4,17,18}. Nevertheless, the origin of adult macrophages and the qualitative and quantitative contributions of HSC and putative non-HSC-derived progenitors are still unclear¹⁹. Here we show in mice that the vast majority of adult tissue-resident macrophages in liver (Kupffer cells), brain (microglia), epidermis (Langerhans cells) and lung (alveolar macrophages) originate from a *Tie2*⁺ (also known as *Tek*) cellular pathway generating *Csf1r*⁺ erythro-myeloid progenitors (EMPs) distinct from HSCs. EMPs develop in the yolk sac at embryonic day (E) 8.5, migrate and colonize the nascent fetal liver before E10.5, and give rise to fetal erythrocytes, macrophages, granulocytes and monocytes until at least E16.5. Subsequently, HSC-derived cells replace erythrocytes, granulocytes and monocytes. Kupffer cells, microglia and Langerhans cells are only marginally replaced in one-year-old mice, whereas alveolar macrophages may be progressively replaced in ageing mice. Our fate-mapping experiments identify, in the fetal liver, a sequence of yolk sac EMP-derived and HSC-derived haematopoiesis, and identify yolk sac EMPs as a common origin for tissue macrophages.

Csf1r-expressing cells in the mouse embryo give rise to tissue-resident macrophages in adult tissues⁴. To identify in the developing embryo the site of origin of *Csf1r*-expressing cells, we performed time course analyses by constitutive (*Csf1r*^{Cre}) and inducible (*Csf1r*^{MerCreMer}) fate-mapping of cells in the yolk sac, head, limbs, caudal region and fetal liver (Fig. 1a and Extended Data Fig. 1). Progenitors, defined as Kit⁺CD45^{lo} (ref. 12) (gate R1 in Fig. 1b), were first detected in *Csf1r*^{Cre}*Rosa26*^{YFP} embryos in the yolk sac from 16–18 somite pairs (sp) stage onwards (E8.5, Fig. 1b, and Extended Data Fig. 1a–c). *Csf1r*^{Cre} YFP⁺ Kit⁺CD45^{lo} cells (gate R2 in Fig. 1b), characterized in Fig. 2 as myeloid cells, were detected in the yolk sac at 20–25 sp (E9, Fig. 1b), and subsequently in the caudal and head regions of the embryo from E9.5, and the fetal liver from E10.5 onwards (Extended Data Fig. 1a–d). To discriminate migration of YFP⁺ cells from *de novo* labelling, we induced YFP expression in *Csf1r*^{MerCreMer}*Rosa26*^{YFP} embryos at E6.5 or E8.5. In embryos pulsed at E6.5, YFP⁺ cells were not detected (Extended Data Fig. 2a, b). When pulsed at E8.5, *Csf1r*^{MerCreMer} YFP⁺ Kit⁺CD45^{lo} progenitors were detected between E9.5–11.5 in the yolk sac, and in the fetal liver from E10.5 (Fig. 1c, d). In the fetal liver, numbers of YFP⁺ Kit⁺CD45^{lo} progenitors increased threefold from E10.5 to E11.5, at which time they were 25-fold more numerous in the fetal liver than in the yolk sac (Fig. 1d). At E8.5, all YS *Csf1r*^{Cre} YFP⁺ Kit⁺CD45^{lo} progenitors expressed AA4.1, an antigen expressed on early haematopoietic progenitors¹² (Extended Data Fig. 1e). *Csf1r*^{MerCreMer} YFP⁺ AA4.1⁺ Kit⁺CD45^{lo} cells were also

present in the yolk sac from E9.5 to E10.5, and in the fetal liver from E10.5 (Fig. 1e). These progenitors were undetectable at E10.5 in the aorta-gonado-mesonephros (AGM) region (Fig. 1e), indicating they do not originate within the embryo proper.

Together, these fate-mapping experiments demonstrate that yolk-sac-derived progenitors colonize the liver primordium, as proposed earlier^{8,20,21}, and their expression of AA4.1 suggests that they represent erythro-myeloid progenitors (EMPs)¹². In *in vitro* colony-forming assays, the AA4.1⁺ population contained most of the total E9 yolk sac colony-forming-units-culture (CFU-C 266 ± 137 vs 296 ± 75, mean ± standard deviation (s.d.)). Frequencies and distributions of different CFU-C, that is, erythroid (E)/megakaryocyte (Mk) (E/Mk), granulocyte/macrophage (G/M), and G,M,E, and/or Mk (Mix) potential, were comparable between overall AA4.1⁺ and *Csf1r*^{Cre} YFP⁺ AA4.1⁺ progenitors (Fig. 2a, Extended Data Fig. 3). Moreover, in the E12.5 fetal liver, the CFU potential of overall AA4.1⁺ and *Csf1r*^{MerCreMer} YFP⁺ AA4.1⁺ cells was comparable to the yolk sac progenitors (Fig. 2a).

These results indicated that yolk-sac-derived, E8.5-labelled YFP⁺ AA4.1⁺ Kit⁺CD45^{lo} progenitors have erythroid and myeloid potential in yolk sac and fetal liver. Next, we investigated by fate-mapping their contribution to fetal liver haematopoiesis *in vivo*. *Csf1r*^{Cre} YFP⁺ and *Csf1r*^{MerCreMer} YFP⁺ F4/80^{bright} fetal macrophages were first detected among Kit⁺CD45⁺ (R2 in Fig. 1b) at E10.5 in the yolk sac, liver, head and forelimbs (Fig. 2b, Extended Data Fig. 4a–c). In addition, the fetal liver from E12.5 to E16.5 contained *Csf1r*^{MerCreMer} YFP⁺ monocytes and granulocytes (Fig. 2c). The fetal liver also contained *Csf1r*^{MerCreMer} YFP⁺ red blood cells from E11.5 until at least E14.5 (Fig. 2d, Extended Data Fig. 4d). Red blood cells were not labelled before E11.5, indicating that, in contrast to yolk-sac-derived erythrocytes in the fetal liver, primitive erythrocytes in the yolk sac did not arise from *Csf1r*-expressing cells. Collectively, yolk-sac-derived *Csf1r*⁺ progenitors contribute to fetal liver haematopoiesis by giving rise to F4/80^{bright} macrophages, monocytes, granulocytes and red blood cells.

We next investigated the transition from yolk-sac-derived to HSC-derived haematopoiesis. To trace the latter, we used *Flt3*^{Cre} which labels fetal and adult HSC-derived multipotent haematopoietic progenitors²², and their progeny (Extended Data Fig. 5). We compared progeny of yolk-sac-derived progenitors in *Csf1r*^{MerCreMer} mice to progeny of HSCs in *Flt3*^{Cre} mice. In the fetal liver from E14.5 to E18.5, the progenies of *Csf1r*⁺ and *Flt3*⁺ precursors were distinct but complemented each other (Fig. 3a). At E14.5, yolk-sac-derived CD45⁺ populations included Kit⁺ progenitors, F4/80^{bright} macrophages, and CD11b^{hi}Gr1⁺ monocytes/granulocytes (Fig. 3a, Extended Data Fig. 6a). Of note, monocytes/granulocytes were present in *Myb*-deficient fetal liver (Fig. 3a). *Csf1r*^{MerCreMer} YFP⁺ macrophages remained detectable throughout fetal development, and were not replaced by *Flt3*^{Cre} YFP⁺ cells. However, yolk-sac-derived Kit⁺ cells and myeloid cells were no longer detectable by E16.5 and E18.5, respectively (Fig. 3a). In contrast, *Flt3*^{Cre} YFP⁺ Kit⁺

¹Centre for Molecular and Cellular Biology of Inflammation (CMCBI), King's College London, London SE1 1UL, UK. ²Division of Cellular Immunology, German Cancer Research Center (DKFZ), D-69120 Heidelberg, Germany. ³MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital University of Oxford, Oxford OX3 9DS, UK.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

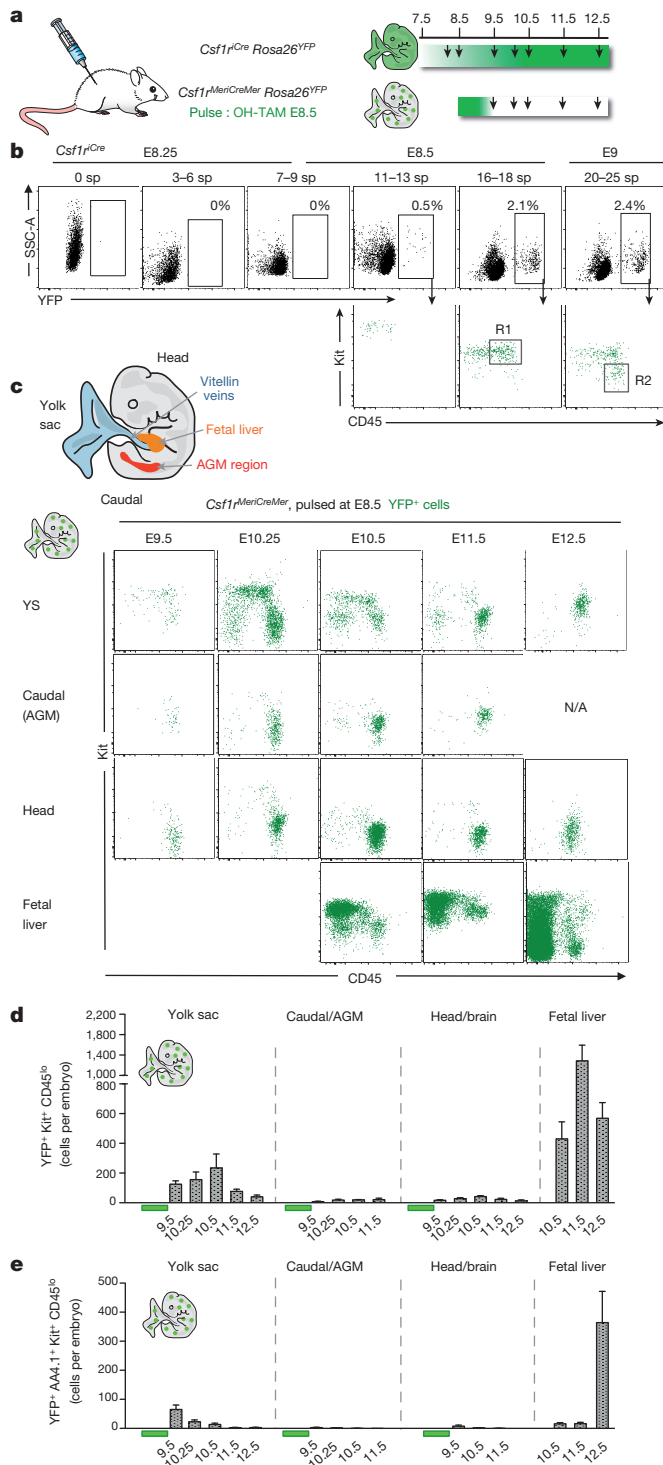


Figure 1 | E8.5 *Csf1r*⁺ progenitors originate in the yolk sac and expand in the fetal liver. **a**, Fate-mapping analysis of *Csf1r*-expressing cells. Arrows indicate time points for analysis, and green shades the genetic labelling period. **b**, YFP expression on live cells from *Csf1r^{Cre} Rosa26^{YFP}* yolk sac (YS), separated by somite pairs between E8.25 and E9 (0 sp, n = 3; 3–6 sp, n = 3; 7–9 sp, n = 5; 11–13 sp, n = 3; 16–18 sp, n = 4; 20–25 sp, n = 4) (upper panels), and Kit and CD45 expression on YFP⁺ cells (lower panels). R1 indicates Kit⁺CD45^{lo}, and R2 indicates Kit⁺CD45^{hi} cells. **c**, Schematic representation of sites analysed in mouse embryos: YS, AGM region, fetal liver and head. Kit and CD45 phenotype of YFP⁺ cells from *Csf1r^{MerCreMer} Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5 (E9.5, n = 3; E10.25, n = 3; E10.5, n = 4; E11.25, n = 4; E12.5, n = 9). **d**, Number of YFP⁺Kit⁺CD45^{lo} cells (R1 in panel b) per organ or region (mean ± s.e.m.) in *Csf1r^{MerCreMer} Rosa26^{YFP}* embryos pulsed at E8.5 (Source Data Table for Fig. 1). **e**, Number of YFP⁺AA4.1⁺Kit⁺CD45^{lo} cells per embryonic region and time points (mean ± s.e.m.) in *Csf1r^{MerCreMer} Rosa26^{YFP}* embryos pulsed at E8.5 (Source Data Table for Fig. 1). See also Supplementary Table 1 and Extended Data Figs 1 and 2.

Altogether, these data indicate that *Flt3^{Cre} YFP⁺ Kit⁺* progenitors and monocytes account for only minor fractions of microglia, Kupffer cells, alveolar macrophages and Langerhans cells in young adults. To investigate whether the presence of these adult *Flt3^{Cre} YFP⁺ F4/80^{bright}* macrophages corresponds to their HSC origin, we performed non-myeloablative transplantations of YFP⁺ long-term-HSCs (LT-HSCs) from adult wild-type bone marrow into *Rag2^{-/-} γc^{-/-} Kit^{W/W^v}* recipients²³ (Extended Data Fig. 7). Eight weeks after transplantation, the vast majority of HSCs, myeloid progenitors, monocytes and F4/80^{lo} tissue myeloid cells in the recipients were of donor HSC origin. In contrast, only 7% of F4/80^{bright} macrophages in spleen, 2% in liver, 5% in lung, 13% in pancreas, 2% in epidermis and 0% in the brain were donor-derived. Thus, recruitment of HSC-derived precursors is not a major mechanism for the maintenance of F4/80^{bright} macrophages in these tissues.

Collectively, these findings reveal that the transition from yolk-sac- to HSC-derived haematopoiesis occurs late in fetal development for monocytes (E14.5) and granulocytes (E16.5), and suggest that HSC-derived progenitors only marginally replace yolk-sac-derived microglia in the brain, Kupffer cells in the liver, Langerhans cells in the epidermis, although alveolar macrophages and brain CD45⁺F4/80⁺ macrophages may undergo progressive replacement with age.

Labelling efficiency of most tissue-resident macrophage populations in adult *Csf1r^{MerCreMer} Rosa26^{YFP}* mice pulse-labelled with 4-hydroxytamoxifen (OH-TAM) at E8.5 was low^{4,24}. The strength of most genetic pulse-labelling systems is that they allow fate-mapping of cells during a specific time window, however, a weakness is the commonly incomplete labelling which could explain why a large fraction of tissue-resident macrophages remained unlabelled. Hence, based on these data we cannot formally exclude a fetal HSC origin of the unlabelled cells as suggested by others based on transfer of fetal precursors^{24–26}.

We thus made use of a newly generated inducible *Cre* knock-in mouse (*Tie2^{MerCreMer}*) to track haematopoietic output from haematopoietic progenitors and HSCs *in situ* (Busch K. *et al.*, submitted). *Tie2* (also known as *Tek*) is expressed in endothelial cells, yolk sac progenitors, aorta-gonado-mesonephros region, fetal liver and adult HSCs³. We assessed the time window at which *Tie2⁺* cells contributed to emerging HSCs and macrophages by injecting tamoxifen at different time points (Fig. 4a, Extended Data Figs 8–10). Fetal liver E12.5 and E15.5 LT-HSCs were labelled efficiently in *Tie2^{MerCreMer}* embryos pulsed at E6.5, E7.5 or E10.5 (Fig. 4b, Extended Data Figs 8 and 9). Yolk sac E9.5 Kit⁺CD45^{lo} progenitors were also labelled in *Tie2^{MerCreMer} Rosa26^{YFP}* embryos pulsed at E7.5 (Extended Data Fig. 10). Interestingly, fetal liver cells with a megakaryocyte-erythrocyte progenitor (MEP) phenotype, and F4/80^{bright} macrophages in yolk sac, brain, and fetal liver were labelled with high efficiency (60%) in embryos pulsed at E6.5 and E7.5, but not in embryos pulsed at E10.5 (Fig. 4b, Extended Data Fig. 8). These fate-mapping experiments directly demonstrate that E12.5 and E15.5 fetal macrophages originate from cells that express *Tie2⁺* as early as E6.5

cells, and CD11b^{hi}Gr1⁺ granulocytes/ monocytes increased in numbers between E14.5 and E18.5 (Fig. 3a–c). The progenies of *Csf1r*⁺ progenitors and *Flt3*⁺ progenitors were also distinct during development in the lung and skin (Extended Data Fig. 6b, c). Quantitative analyses in fetal and adult tissues indicated that *Flt3^{Cre} YFP* labelling of Kit⁺ progenitors preceded that of monocytes/granulocytes, with 80% of progenitors labelled at E18.5, and 80% of monocytes at postnatal day 8 (P8) (Fig. 3b). In contrast, *Flt3^{Cre} YFP* labelling plateaued at 14% for adult liver F4/80^{bright} Kupffer cells, at 2% for CD45^{lo} brain microglia, and 30% for epidermal Langerhans cells up to one year of life (Fig. 3b). In contrast, *Flt3^{Cre} YFP* labelling of CD45⁺ F4/80⁺ brain macrophages, and lung alveolar macrophages was 16% in 12-week-old adults but increased progressively over time to reach 40% in one-year-old mice (Fig. 3b).

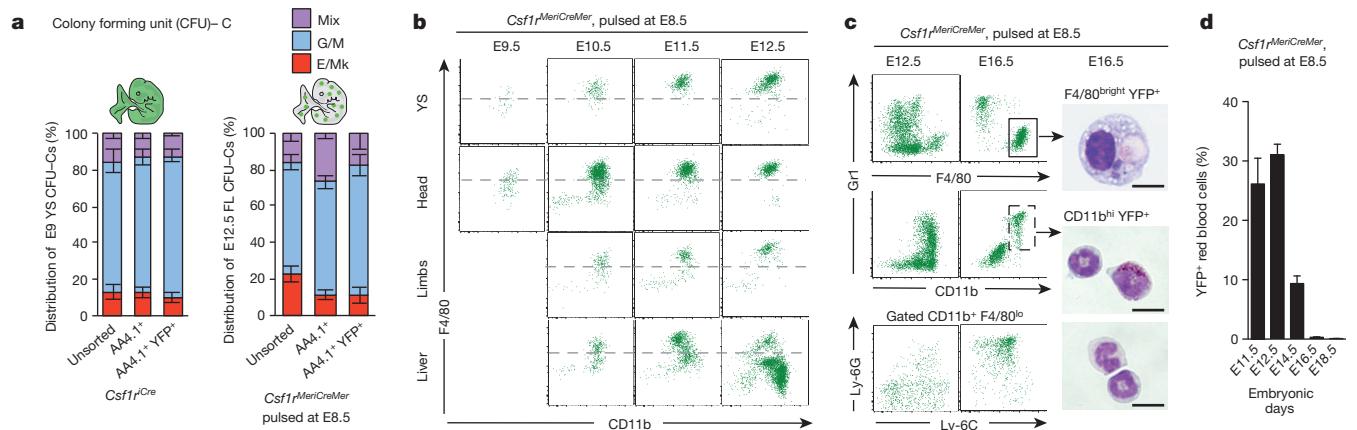


Figure 2 | E8.5 *Csf1r*⁺ progenitors differentiate into myeloid cells and red blood cells in the fetal liver. **a**, Distribution of mixed (Mix), G/M, and E/Mk CFU-C from unsorted, AA4.1⁺ Kit⁺ CD45^{lo} and YFP⁺ AA4.1⁺ Kit⁺ CD45^{lo} cells from E9 yolk sac from *Csf1r*^{iCre} *Rosa26*^{YFP} and E12.5 fetal liver from *Csf1r*^{MeriCreMer} *Rosa26*^{YFP} embryos pulsed with OH-TAM at E8.5 (three independent experiments each). CFU-C erythroid and/or megakaryocyte (E/Mk); CFU-granulocyte and/or monocyte/macrophage (G/M); CFU-C mix, at least three of the following: G, E, M and Mk. See also Extended Data Fig. 3. **b**, F4/80 and CD11b expression on YFP⁺ CD45⁺ from yolk sac, head (brain for E11.5), limbs and liver of *Csf1r*^{MeriCreMer} *Rosa26*^{YFP} embryos pulsed with OH-TAM at E8.5, and analysed on E9.5 (*n* = 3), E10.5 (*n* = 4), E11.5 (*n* = 4), and E12.5 (*n* = 9); Dashed lines represent FMO (fluorescence minus one)

control. See also Extended Data Fig. 4. **c**, F4/80, CD11b, Gr1, Ly-6G and Ly-6C expression in fetal liver CD45⁺ YFP⁺ cells from *Csf1r*^{MeriCreMer} *Rosa26*^{YFP} embryos pulse-labelled at E8.5, and analysed on E12.5 (*n* = 9) and E16.5 (*n* = 14). May-Grünwald-Giemsa stained cytopsin preparations of fetal liver YFP⁺ F4/80^{bright} and YFP⁺ CD11b^{hi} cells sorted from E16.5 *Csf1r*^{MeriCreMer} *Rosa26*^{YFP} embryos pulsed with OH-TAM at E8.5. See also Extended Data Fig. 6a for sorted cells from E14.5 *Csf1r*^{MeriCreMer} embryos. Scale bar, 10 μ m. **d**, YFP labelling efficiency (%) among red blood cells in fetal liver from *Csf1r*^{MeriCreMer} *Rosa26*^{YFP} embryos pulsed with OH-TAM at E8.5, mean \pm s.e.m. (E11.5, *n* = 4; E12.5, *n* = 9; E14.5, *n* = 5; E16.5, *n* = 11; E18.5, *n* = 4; see Source Data Table Fig. 3).

and, importantly, before E9.5, and strongly support the notion that fetal liver erythro-myeloid progenitors, and all fetal tissue macrophages up to E15.5 are of yolk sac origin.

In adult mice pulsed at embryonic stages (E7.5, or E8.5, or E9.5 or E10.5), bone marrow HSC-derived progenitors, peripheral cells (T and B cells, and granulocytes) in the spleen, and CD11b^{hi} F4/80^{lo} myeloid cells

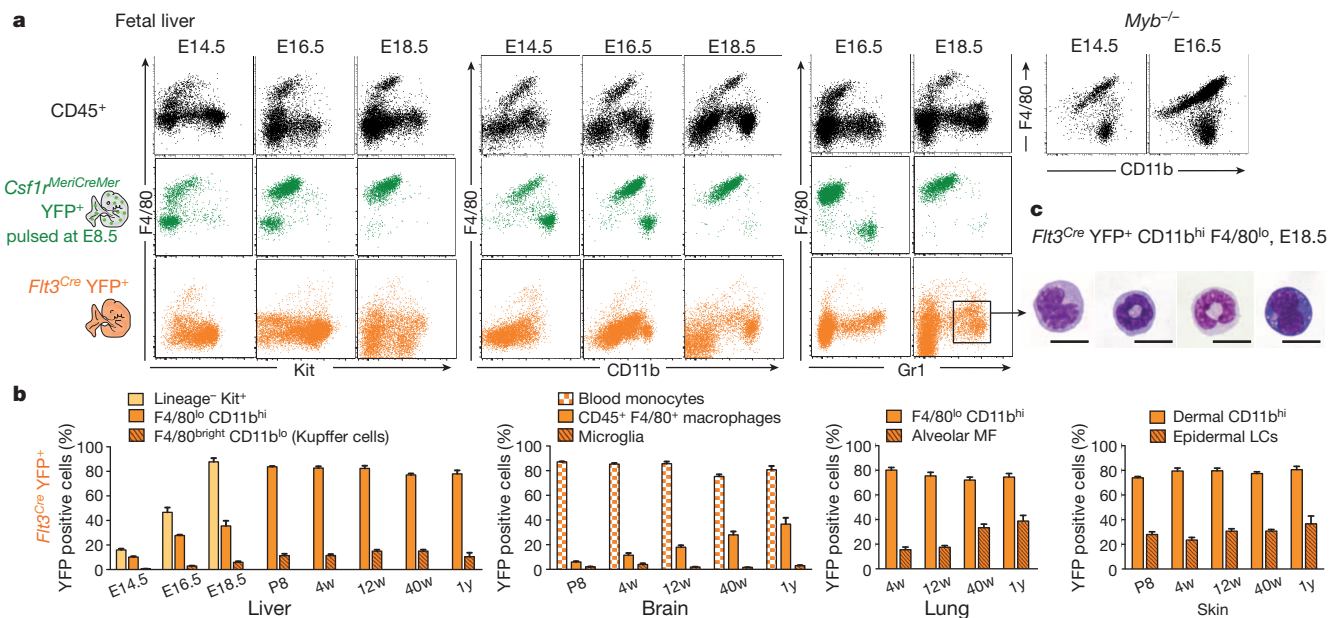


Figure 3 | Fetal liver HSC-derived *Flt3*⁺ progenitors give rise to monocytes and granulocytes in late embryos and adults but do not replace yolk-sac-derived macrophages. **a**, F4/80, Kit, CD11b and Gr1 expression on total CD45⁺ cells (black) and YFP⁺ CD45⁺ cells from *Csf1r*^{MeriCreMer} *Rosa26*^{YFP} embryos pulsed at E8.5 (green) in the fetal liver at the indicated days of embryonic development (E14.5, *n* = 5; E16.5, *n* = 10; E18.5, *n* = 9). F4/80, Kit, CD11b and Gr1 expression on YFP⁺ CD45⁺ cells from *Flt3*^{Cre} *Rosa26*^{YFP} embryos (orange) (E14.5, *n* = 7; E16.5, *n* = 6; E18.5, *n* = 6). F4/80 and CD11b expression on CD45⁺ cells in *Myb*^{-/-} embryos (E14.5, *n* = 4; E16.5, *n* = 7). **b**, YFP labelling efficiency in Kit⁺ Lin⁻ cells, CD11b^{hi} F4/80^{lo} cells (characterized in Extended Data Fig. 5) and F4/80^{bright} macrophages (Kupffer cells in adults) in fetal and adult *Flt3*^{Cre} *Rosa26*^{YFP} liver (first panel on the left).

YFP labelling efficiency in blood monocytes, brain microglia (CD45^{lo} F4/80⁺) and CD45⁺ F4/80⁺ brain macrophages in *Flt3*^{Cre} *Rosa26*^{YFP} pups and mice (second panel). YFP labelling efficiency in alveolar macrophages (F4/80^{bright} Siglec-F⁺ CD11b⁺) and F4/80^{lo} CD11b^{hi} myeloid cells in *Flt3*^{Cre} *Rosa26*^{YFP} lungs (third panel). YFP labelling efficiency in epidermal Langerhans cells (LCs) and dermal CD11b^{hi} (MHC II⁺ EpCAM⁺) myeloid cells in *Flt3*^{Cre} *Rosa26*^{YFP} skin (fourth panel, see Extended Data Fig. 6b, c). Mean \pm s.e.m.; P8, *n* = 3; 4-week-old, *n* = 6; 12-week-old, *n* = 11–14; 40-week-old, *n* = 7; 1-year-old, *n* = 3, see Source Data table Fig. 3. w, week; y, year. **c**, Representative images of May-Grünwald-Giemsa stained cytopsin preparations of YFP⁺ CD11b^{hi} F4/80^{lo} cells sorted from E18.5 *Flt3*^{Cre} *Rosa26*^{YFP} fetal liver. Scale bar, 10 μ m.

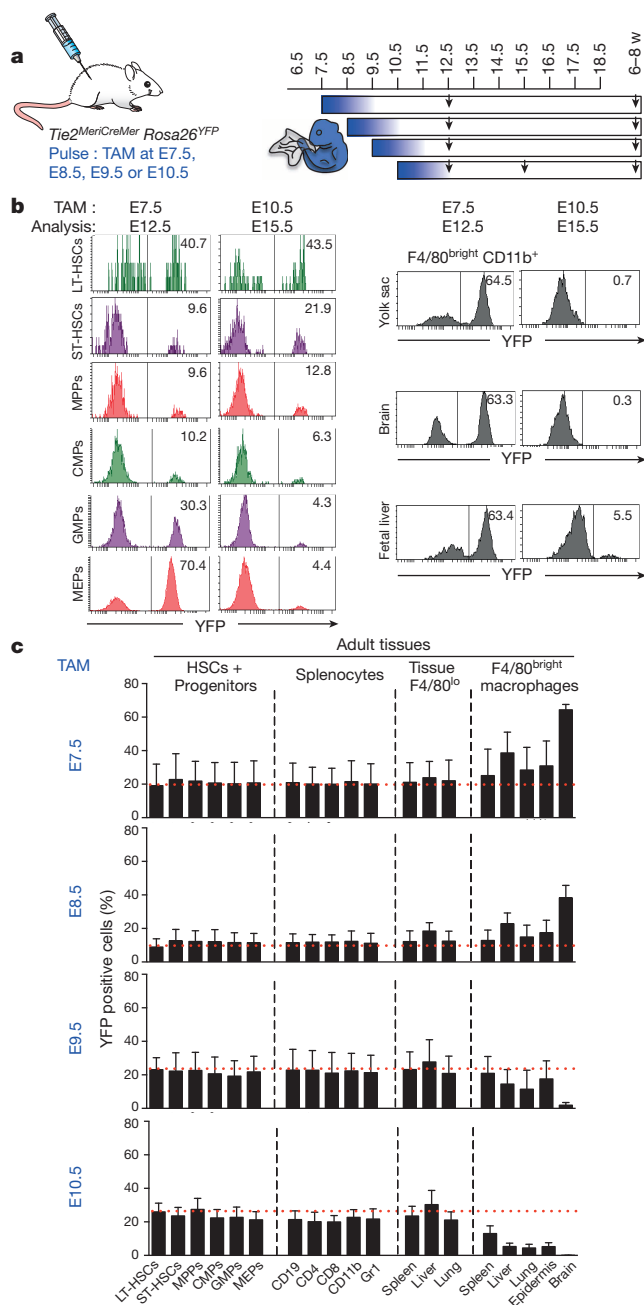


Figure 4 | Fetal macrophages and adult tissue-resident macrophages originate from *Tie2*-expressing progenitors before E10.5. **a**, Fate-mapping analysis of *Tie2*-expressing cells after tamoxifen (TAM) administration at E7.5, or E8.5, or E9.5 or E10.5. Arrows indicate time points for analysis. **b**, Flow cytometric analysis of fetal liver long-term or short-term haematopoietic stem cells (LT-HSCs, ST-HSCs), multipotent progenitors (MPPs), common myeloid progenitors (CMPs), granulocyte-monocyte progenitors (GMPs), megakaryocyte-erythrocyte progenitors (MEPs) (left panel) and of fetal macrophages (right panel) in the yolk sac, brain, and fetal liver. Time points of labelling (E7.5 ($n = 7$); E10.5 ($n = 7$)) and analysis are indicated, and for each experiment one representative analysis is shown. See Extended Data Fig. 8 for quantitative analysis. **c**, Frequencies of labelled HSCs and progenitor cells, splenocytes, and F4/80^{lo}CD11b^{hi} myeloid cells and F4/80^{bright} resident macrophages in spleen, liver, lung, epidermis and brain were analysed (mean \pm s.d., see Source Data Table Fig. 4) from 6–8-week-old *Tie2*^{MeriCreMer} animals pulse-labelled at E7.5 ($n = 4$), E8.5 ($n = 4$), E9.5 ($n = 4$) or E10.5 ($n = 6$).

in peripheral tissues (spleen, liver and lung) were homogeneously labelled at frequencies comparable to HSC labelling, consistent with their adult HSC origin (Fig. 4c). In contrast, YFP labelling frequencies of adult

tissue-resident macrophages were maximal in animals pulse-labelled at E7.5, declined at later time points and were minimal when labelled at E10.5 (Fig. 4c). The fact that adult HSCs are disconnected from resident macrophages is further underscored by the finding that resident macrophages in mice pulsed at E7.5 were labelled at higher frequencies than adult HSCs, that is, labelling efficiency did not equilibrate with mouse development. In summary, these inducible temporal analyses demonstrate that although both macrophages and HSCs originate from progenitors expressing *Tie2* as early as E6.5, adult tissue-resident macrophages in the brain (microglia), liver (Kupffer cells), lung (alveolar macrophages), skin (Langerhans cells), and (to some extent) spleen (F4/80^{bright} macrophages) develop almost exclusively from an *Tie2*-expressing progenitor pathway distinct from HSCs. These data are consistent with results from *Csf1r*^{MeriCreMer} pulse-labelling experiments (see Figs 1 and 2), with our earlier observation that resident macrophages are independent of the transcription factor *Myb*⁴, and complement our data obtained in *Flt3*^{Cre}*Rosa26*^{YFP} mice (see Fig. 3).

This study demonstrates that *Myb*-independent tissue-resident macrophages⁴ originate from yolk-sac-derived EMPs, characterized by expression of *Csf1r* from E8.5 (16–18 somites). The data do not distinguish whether resident macrophages originate from erythro-myeloid, granulocyte-macrophage, or macrophage-only-progenitors because these potentials coexist within the yolk-sac-derived EMP population.

We also provide strong *in vivo* evidence for engraftment of yolk-sac-derived EMPs in the early fetal liver. These cells substantially contribute to the first wave of fetal liver haematopoiesis, followed later by bona fide fetal liver HSC-derived haematopoiesis^{8,19,21}. Conclusions from recent studies that Langerhans cells and alveolar macrophages are not of yolk sac origin based on transfer of fetal precursors^{24–26} should be interpreted in light of our findings that yolk sac EMPs expand in the fetal liver and are the main source for tissue-resident macrophages.

Under steady-state conditions, yolk-sac-derived macrophages are only marginally replaced by HSC-derived cells in the brain, liver and epidermis. It is remarkable that macrophages of yolk sac origin persist in functionally very distinct tissues, suggesting that the origin is more deterministic of the life span than the tissue location. However, some yolk-sac-derived macrophages can undergo replacement in older mice, as for lung alveolar macrophages. In a third group, exemplified by gut-associated macrophages²⁷, yolk-sac-derived macrophages are replaced by HSC-derived macrophages in the first weeks of post-natal life. The mechanisms responsible for the maintenance of yolk-sac-derived macrophages in certain adult tissues require further investigation. Although yolk-sac- and HSC-derived macrophages can co-exist in the same environment, and their balance be perturbed by pathology, the contributions of these developmentally distinct macrophage populations to homeostasis and inflammation remain to be characterized.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 April; accepted 21 October 2014.

Published online 3 December 2014.

- Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
- Smith, L. G., Weissman, I. L. & Heimfeld, S. Clonal analysis of hematopoietic stem-cell differentiation *in vivo*. *Proc. Natl Acad. Sci. USA* **88**, 2788–2792 (1991).
- Cumano, A. & Godin, I. Ontogeny of the hematopoietic system. *Annu. Rev. Immunol.* **25**, 745–785 (2007).
- Schulz, C. *et al.* A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* **336**, 86–90 (2012).
- Yona, S. *et al.* Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* **38**, 79–91 (2013).
- Hashimoto, D. *et al.* Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity* **38**, 792–804 (2013).
- Ajami, B., Bennett, J. L., Krieger, C., Tetzlaff, W. & Rossi, F. M. Local self-renewal can sustain CNS microglia maintenance and function throughout adult life. *Nature Neurosci.* **10**, 1538–1543 (2007).

8. Palis, J., Robertson, S., Kennedy, M., Wall, C. & Keller, G. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126**, 5073–5084 (1999).
9. Bertrand, J. Y. *et al.* Three pathways to mature macrophages in the early mouse yolk sac. *Blood* **106**, 3004–3011 (2005).
10. Lux, C. T. *et al.* All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac. *Blood* **111**, 3435–3438 (2008).
11. Kierdorf, K. *et al.* Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nature Neurosci.* **16**, 273–280 (2013).
12. Bertrand, J. Y. *et al.* Characterization of purified intraembryonic hematopoietic stem cells as a tool to define their site of origin. *Proc. Natl Acad. Sci. USA* **102**, 134–139 (2005).
13. Moore, M. A. & Metcalf, D. Ontogeny of the haematopoietic system: yolk sac origin of *in vivo* and *in vitro* colony forming cells in the developing mouse embryo. *Br. J. Haematol.* **18**, 279–296 (1970).
14. Herbolme, P., Thisse, B. & Thisse, C. Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development* **126**, 3735–3745 (1999).
15. Takahashi, K., Yamamura, F. & Naito, M. Differentiation, maturation, and proliferation of macrophages in the mouse yolk sac: a light-microscopic, enzyme-cytochemical, immunohistochemical, and ultrastructural study. *J. Leukoc. Biol.* **45**, 87–96 (1989).
16. Sumner, R., Crawford, A., Mucenski, M. & Frampton, J. Initiation of adult myelopoiesis can occur in the absence of c-Myb whereas subsequent development is strictly dependent on the transcription factor. *Oncogene* **19**, 3335–3342 (2000).
17. Alliot, F., Godin, I. & Pessac, B. Microglia derive from progenitors, originating from the yolk sac, and which proliferate in the brain. *Brain Res. Dev. Brain Res.* **117**, 145–152 (1999).
18. Ginhoux, F. *et al.* Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* **330**, 841–845 (2010).
19. Frame, J. M., McGrath, K. E. & Palis, J. Erythro-myeloid progenitors: “definitive” hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells Mol. Dis.* **51**, 220–225 (2013).
20. Dzierzak, E. & Medvinsky, A. Mouse embryonic hematopoiesis. *Trends Genet.* **11**, 359–366 (1995).
21. Kieusseian, A., Brunet de la Grange, P., Burlen-Defranoux, O., Godin, I. & Cumano, A. Immature hematopoietic stem cells undergo maturation in the fetal liver. *Development* **139**, 3521–3530 (2012).
22. Christensen, J. L. & Weissman, I. L. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc. Natl Acad. Sci. USA* **98**, 14541–14546 (2001).
23. Waskow, C. *et al.* Hematopoietic stem cell transplantation without irradiation. *Nature Methods* **6**, 267–269 (2009).
24. Epelman, S. *et al.* Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* **40**, 91–104 (2014).
25. Hoeffel, G. *et al.* Adult Langerhans cells derive predominantly from embryonic fetal liver monocytes with a minor contribution of yolk sac-derived macrophages. *J. Exp. Med.* **209**, 1167–1181 (2012).
26. Guillemin, M. *et al.* Alveolar macrophages develop from fetal monocytes that differentiate into long-lived cells in the first week of life via GM-CSF. *J. Exp. Med.* **210**, 1977–1992 (2013).
27. Bain, C. C. *et al.* Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice. *Nature Immunol.* **15**, 929–937 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors are indebted to J. Pollard, University of Edinburgh for the *Csf1r* reporter strains, J. Frampton, University of Birmingham for the *Myb*-deficient animals, and T. Boehm, Max Planck Institute, Freiburg for the *Flt3^{Cre}* strain. The authors also thank A. McGuigan and the staff of the Biological Service Unit at King's College London, S. Heck and the Biomedical Research Centre at King's Health Partners, S. Woodcock and the staff of the Viapath haematology laboratory in Guy's hospital and S. Schäfer and T. Arnsperger for technical assistance at the German Cancer Research Center. This work was supported by a Wellcome Trust Senior Investigator award (WT101853MA) and ERC Investigator award (2010-StG-261299) from the European Research Council to F.G. and an ERC Investigator award (Advanced Grant 233074), SFB 938 project L, and SFB 873 project B11 to H.-R.R.

Author Contributions E.G.P. and F.G. designed the study and wrote the manuscript. E.G.P., C.S., L.C., H.G. and C.T. performed fate-mapping experiments and E.G.P. and F.G. designed experiments and analysed the data. K.B. and H.-R.R. generated the *Tie2^{MerCreMer}* strain and K.K., K.B. and H.-R.R. designed, performed and analysed fate-mapping experiments. E.A. and E.G.P. performed the CFU assays and E.A., E.G.P., F.G. and M.F.d.B. analysed and interpreted the experimental data. All authors contributed to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.G. (frederic.geissmann@kcl.ac.uk).

METHODS

Animals. *Myb*^{-/-} (ref. 28), *Csf1r*^{MerCreMer} (ref. 29), *Csf1r*^{iCre} (ref. 30), *Flt3*^{Cre} (ref. 31), *Rag2*^{-/-} γ_c ^{-/-}*Kit*^{W/W^v} (ref. 23) and *Rosa26*^{YFP} reporter³² mice have been previously described. *Rag2*^{-/-} γ_c ^{-/-}*Kit*^{W/W^v} were on a mixed genetic background²³, *Csf1r*^{MerCreMer} and *Csf1r*^{iCre} mice were on FVB background, other mice were on C57BL/6 (CD45.2) background.

Csf1r^{MerCreMer} and *Csf1r*^{iCre} mice were generated and provided by J. W. Pollard. *Myb*^{-/-} mice were generated and provided by J. Frampton. *Flt3*^{Cre} mice were generated by C. Bleul and provided by T. Boehm and S. E. Jacobsen. *Rosa26*^{YFP} (B6.129X1-Gt(ROSA)26Sor^{tm1(EYFP)Cos/J}) reporter mice were purchased from The Jackson Laboratory. To achieve tamoxifen-dependent Cre activity in *Tie2*-expressing cells *in vivo* we generated *Tie2*^{MerCreMer} mice. We inserted a tamoxifen inducible codon-improved recombinase (iCre)³³ flanked by two mutated oestrogen receptor sites (MerCreMer)³⁴ into the first exon of the *Tie2* locus. *Tie2*^{MerCreMer} mice were crossed to *Rosa26*^{YFP} mice to fate-map the progeny of *Tie2*-expressing cells.

No randomization method was used and the investigators were blinded to the genotype of the embryos and animals during the experimental procedure. Results are displayed as mean \pm s.e.m. (Figs 1, 2 and 3) or s.d. (Fig. 4, Supplementary Table 1). All experiments included littermate controls and the minimum sample size used was 3. Embryonic development was estimated considering the day of vaginal plug formation as 0.5 days post-coitum (dpc), and staged by developmental criteria⁸. In Figs 1 and 2 and Extended Data Figs 1–4, embryos were included based on their somite number for embryonic days < E11.5, as described in ref. 8. No statistical method was used to predetermine sample size.

All animal procedures were performed in adherence to our project licence issued by the United Kingdom Home Office under the Animals (Scientific Procedures) Act 1986, or by the German regional council at the Regierungspräsidium Karlsruhe, Germany, respectively.

Genotyping. PCR genotyping of *Myb*²⁸, *Csf1r*^{iCre} (ref. 30), *Rag2*^{-/-} γ_c ^{-/-}*Kit*^{W/W^v} (ref. 23), *Csf1r*^{MerCreMer} and *Flt3*^{Cre} mice³⁴ was performed according to protocols described previously. PCR genotyping of *Tie2*^{MerCreMer} will be described elsewhere (Busch K. *et al.*, submitted).

Processing of tissues for flow cytometry. Pregnant females were killed by cervical dislocation or by exposure to CO₂. Embryos ranging from embryonic day (E) 8.25 to E18.5 were removed from the uterus and washed in 4°C phosphate-buffered saline (PBS, Invitrogen). The yolk sac (YS) was harvested from embryos between E8.25 and E12.5. Embryos were exsanguinated through decapitation in PBS 1 × 10 mM EDTA. To obtain single-cell suspensions, organs were incubated in PBS containing 1 mg ml⁻¹ collagenase D (Roche), 100 U ml⁻¹ DNase I (Sigma) and 3% fetal calf serum (FCS, Invitrogen) at 37°C for 30 min.

Adult tissues (P8 to 1 year) were prepared as follows. Blood was collected by cardiac puncture from anaesthetized (isoflurane inhalation) mice. Under terminal anaesthesia, mice were perfused by gentle intracardiac injection of 10 ml prewarmed (37°C) 1 × PBS. The spleen, right liver lobe and right lung lobes were collected and processed for flow cytometry. To obtain single-cell suspensions, organs were incubated for 30 min in PBS containing 1 mg ml⁻¹ Collagenase D (Roche), 100 U ml⁻¹ DNase I (Sigma), 2.4 mg ml⁻¹ of dispase (Invitrogen) and 3% FCS (Invitrogen) at 37°C. Brains from *Tie2*^{MerCreMer} were dissociated and incubated for 60 min at 37°C in HBSS with 0.2 mg ml⁻¹ collagenase D, 20 µg ml⁻¹ dispase I (Roche), and 50 U ml⁻¹ DNase I (Sigma). Brain cells were resuspended in isotonic Percoll (Pharmacia) at a final density of 1.072 g ml⁻¹ in HBSS containing 3% FCS. The suspension was underlayered with Percoll solution at 1.088 g ml⁻¹ and overlaid with additional layers of Percoll (1.06, 1.05 and 1.03 g ml⁻¹). After centrifugation, cells were collected from 1.06 and 1.072 g ml⁻¹ layers. Brains from the other strains were processed as described for the spleen. For collection of Langerhans cells from *Tie2*^{MerCreMer} mice, epidermal sheets were prepared using an epidermis dissociation kit (Miltenyi Biotec). In the other strains, epidermal sheets were separated from the dermis after incubation for 45 min at 37°C in 2.4 mg ml⁻¹ of dispase (Invitrogen) and 3% FCS (Invitrogen) and the epidermis was further digested for 30 min in PBS containing 1 mg ml⁻¹ collagenase D (Roche), 100 U ml⁻¹ DNase I (Sigma), 2.4 mg ml⁻¹ of dispase (Invitrogen) and 3% FCS (Invitrogen) at 37°C.

Flow cytometric analysis of embryonic and adult tissues and cell sorting. Tissues were mechanically dissociated and passed through a 100 µm cell strainer (BD). Red blood cell lysis of fetal liver and adult lung and spleen was performed as described³⁵. Cells were centrifuged at 320g for 7 min, resuspended in 4°C PBS, plated in multi-well round-bottom plates and immunolabelled for FACS analysis. After 15 min incubation with purified anti-CD16/32 (FcγRIII/II) diluted 1/50, or ChromPure mouse IgG whole molecule (Dianova) diluted 1/20 in staining buffer (1 × PBS; 0.5% BSA; 2 mM EDTA), antibody mixes were added and incubated for 30 min. Where appropriate, cells were further incubated with streptavidin conjugates for 20 min. The full list of antibodies used can be found in Supplementary Table 2.

Flow cytometry was performed using a BD Biosciences FACSCanto II flow cytometer or a BD Biosciences LSR Fortessa cell analyser. All data were analysed using FlowJo 9.5 (Tree Star) or FACS Diva software (BD Bioscience).

Fetal liver, skin and lung YFP⁺F4/80^{bright} and YFP⁺CD11b^{hi} cells from E18.5 *Flt3*^{Cre}*Rosa26*^{YFP} embryos and from E14.5 and E16.5 *Csf1r*^{MerCreMer}*Rosa26*^{YFP} embryos pulsed at E8.5 were sorted into FCS-coated tubes using FACSaria II for cytospin preparations.

Pulse labelling of *Csf1r*⁺ and *Tie2*⁺ progenitors. For genetic cell labelling we crossed tamoxifen-inducible *Csf1r*^{MerCreMer} and *Tie2*^{MerCreMer} transgenic mouse strains with *Rosa26*^{YFP} reporter mice. In *Csf1r*^{MerCreMer}*Rosa26*^{YFP} embryos recombination was induced by single injection at E8.5 of 75 µg per g (body weight) of 4-hydroxytamoxifen (Sigma) into pregnant females. The 4-hydroxytamoxifen was supplemented with 37.5 µg per g (body weight) progesterone (Sigma). In *Tie2*^{MerCreMer}*Rosa26*^{YFP} embryos recombination was induced by treatment of pregnant females by gavage at different time points (between E7.5 and E10.5) with a single dose of 2.5 mg tamoxifen (Sigma) and 1.75 mg progesterone (Sigma) to counteract the mixed oestrogen agonist effects of tamoxifen, which can result in fetal abortions.

Continuous labelling of *Csf1r*⁺ progenitors. For fate-mapping analysis of *Csf1r*⁺ precursors, *Csf1r*^{iCre} females were crossed with homozygous *Rosa26*^{YFP} reporter males. Indicated tissues from embryos and adult F1 mice were analysed by flow cytometry.

Fate-mapping of *Flt3*⁺ haematopoietic progenitors. For fate-mapping analysis of *Flt3*⁺ precursors, *Flt3*^{Cre} males (the transgene is located on the Y chromosome) were crossed to homozygous *Rosa26*^{YFP} reporter females. For adult experiments, *Flt3*^{Cre} males were blood phenotyped. Animals with YFP labelling efficiency above 60% in the lymphocytes, monocytes and granulocytes were used for experiments and female littermates were used as Cre-negative controls.

Colony forming assays. Colony-forming-unit-culture (CFU-C) assays were performed using Methocult M3434 (Stem Cell Technologies) as described in ref. 36. Embryos were collected and dissected in PBS (Gibco, Invitrogen) supplemented with 10% FCS (batch tested and obtained from Gibco), 50 U ml⁻¹ penicillin, and 50 µg ml⁻¹ streptomycin (Cambrex Corporation). E9 embryos were staged by somite counting. E9 yolk sac and E12 fetal livers were each pooled and incubated for 30 min at 37°C in PBS supplemented with 10% FCS, 50 U ml⁻¹ penicillin, 50 µg ml⁻¹ streptomycin, 1 mg ml⁻¹ collagenase D (Roche) and 100 U ml⁻¹ DNase I (Sigma), and dissociated by pipetting. Suspensions were washed, and viable cells were counted on the basis of trypan blue (Sigma) exclusion using a Kova hemocytometer slide.

AA4.1⁺ progenitors were isolated by flow cytometry using FACSaria II or FACSaria III. Labelling of cells was performed as described above using the following antibodies: CD45-APC-Cy7, Kit-PE and AA4.1-APC (Supplementary Table 2), and live cells were gated on the basis of Hoechst 33258 exclusion. Cells were collected into FCS-coated tubes and recounted before plating where possible. Gates were defined using unstained, single stained and fluorescence minus one (FMO) stained cells.

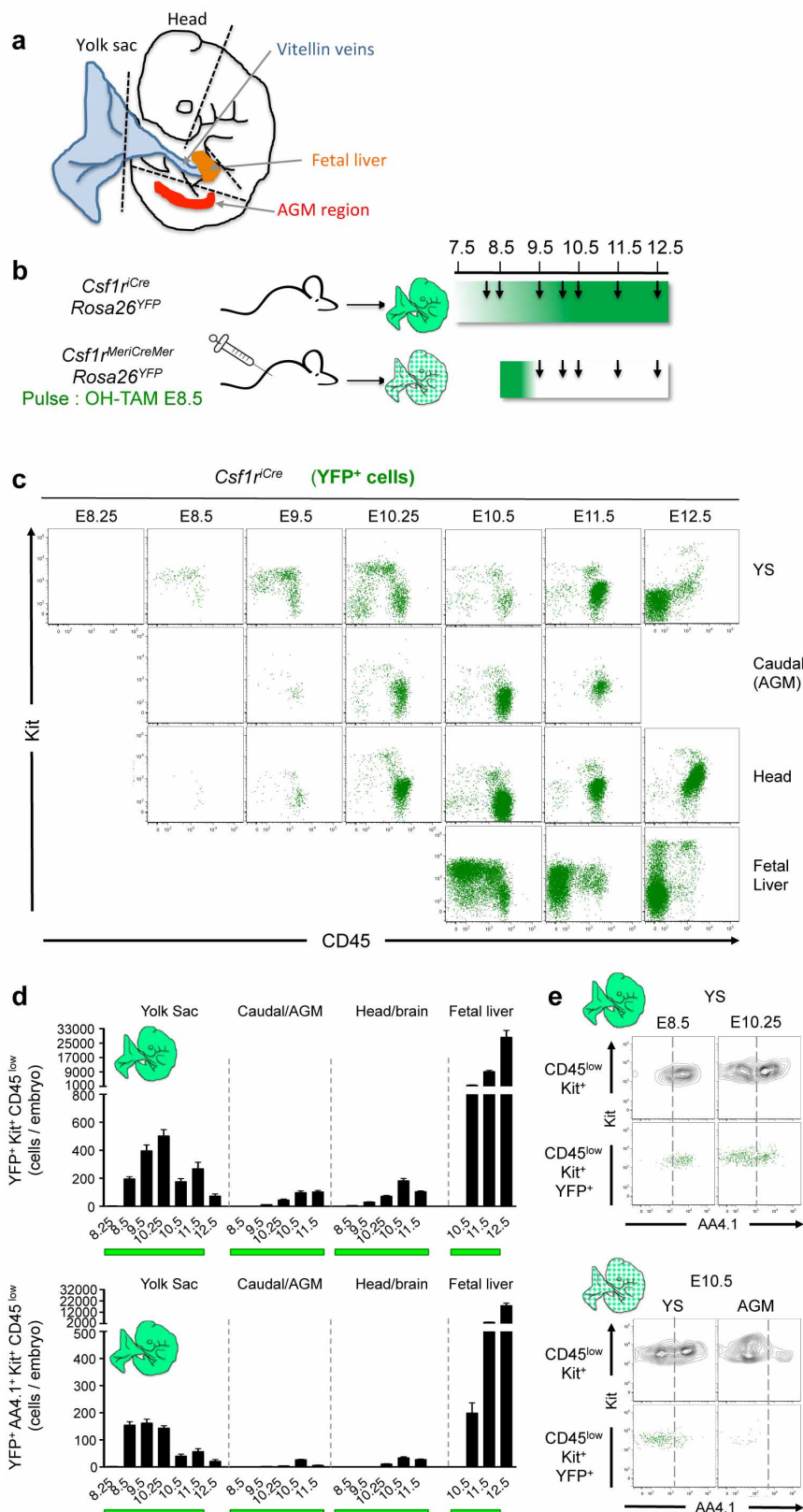
Cells were plated in duplicate in 35-mm culture dishes according to manufacturer's instructions. Cultures were grown at 37°C with 5% CO₂ with colonies scored after 10 days.

Colonies were picked and washed once with phosphate-buffered saline (PBS; Gibco, Invitrogen) supplemented with 10% fetal calf serum (FCS; batch tested and obtained from Gibco). Cytospin preparations were stained with May-Grünwald-Giemsa method for morphological inspection of colonies (see below).

Morphological analysis of sorted cells and colonies. Cytospin preparations were performed using a Cytospin 3 (Thermo Shandon) by centrifuging (i) cells from colonies at 400 r.p.m. for 4 min (medium acceleration) or (ii) sorted cells at 500 r.p.m. for 10 min (low acceleration). Slides were air-dried for at least 30 min, and fixed for 5 min in methanol. Methanol-fixed cytospin preparations were manually stained in 50% May-Grünwald solution for 5 min, 14% Giemsa for 15 min, washed with Sorensons buffered distilled water (pH 6.8) for 5 min and rinsed with Sorensons buffered distilled water (pH 6.8). After air-drying, slides were mounted with Entellan New (Merck) and representative pictures were taken using a Nikon eclipse E6000 microscope with a Nikon Plan Fluor 60×/1.40 NA oil DIC H objective and NIS-elements BR2.30 software (Nikon).

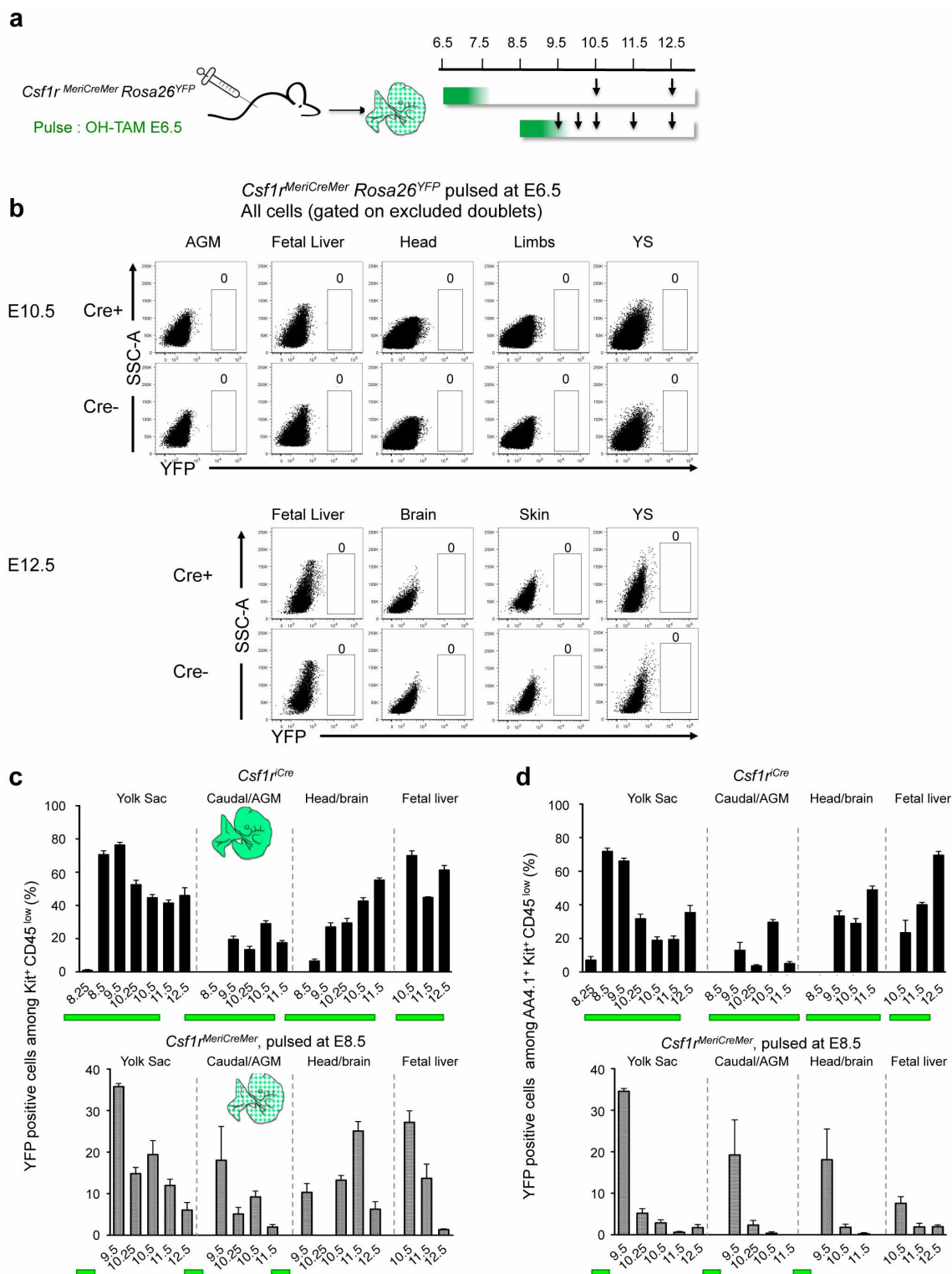
Transplantation of HSCs without irradiation. HSC transplantation in non-irradiated *Rag2*^{-/-} γ_c ^{-/-}*Kit*^{W/W^v} mice was performed as described previously²³. In brief, approximately 1000 LT-HSCs (lin⁻Sca-1⁺Kit⁺CD150⁺CD48⁻) isolated from the bone marrow of *panRosa*^{YFP} mice, which carry a constitutively active YFP reporter allele, were injected into *Rag2*^{-/-} γ_c ^{-/-}*Kit*^{W/W^v} mice. Recipients were analysed 2 months after transplantation for donor/host chimaerism in blood, spleen, lung, liver, pancreas, brain and epidermis. To test the functionality of E12.5 phenotypic LT-HSCs, 10 YFP⁺ LSK CD150⁺CD48⁻ (phenotypic LT-HSCs) from *Tie2*^{MerCreMer}*Rosa26*^{YFP} pulsed at E7.5 were transplanted into *Rag2*^{-/-} γ_c ^{-/-}*Kit*^{W/W^v} mice and blood lineages were analysed 16 weeks after.

28. Mucenski, M. L. *et al.* A functional *c-myb* gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65**, 677–689 (1991).
29. Qian, B. Z. *et al.* CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* **475**, 222–225 (2011).
30. Deng, L. *et al.* A novel mouse model of inflammatory bowel disease links mammalian target of rapamycin-dependent hyperproliferation of colonic epithelium to inflammation-associated tumorigenesis. *Am. J. Pathol.* **176**, 952–967 (2010).
31. Benz, C., Martins, V. C., Radtke, F. & Bleul, C. C. The stream of precursors that colonizes the thymus proceeds selectively through the early T lineage precursor stage of T cell development. *J. Exp. Med.* **205**, 1187–1199 (2008).
32. Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of *EYFP* and *EGFP* into the *ROSA26* locus. *BMC Dev. Biol.* **1**, 4 (2001).
33. Shimshek, D. R. *et al.* Codon-improved Cre recombinase (iCre) expression in the mouse. *Genesis* **32**, 19–26 (2002).
34. Zhang, Y. *et al.* Inducible site-directed recombination in mouse embryonic stem cells. *Nucleic Acids Res.* **24**, 543–548 (1996).
35. Auffray, C. *et al.* CX3CR1⁺ CD115⁺ CD135⁺ common macrophage/DC precursors and the role of CX3CR1 in their response to inflammation. *J. Exp. Med.* **206**, 595–606 (2009).
36. Swiers, G. *et al.* Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nature Commun.* **4**, 2924 (2013).



Extended Data Figure 1 | Analysis of *Csf1r* reporter expression in fetal progenitor cells in *Csf1r^{flCre} Rosa26^{YFP}*. **a**, Schematic representation of the different haematopoietic and non-haematopoietic sites dissected in the mouse embryos: yolk sac (YS), aorta-gonado-mesonephros (AGM) region, fetal liver and head. **b**, Experimental design for fate-mapping analysis of *Csf1r*-expressing cells. Arrows indicate analysed time points **c**, Kit and CD45 expression on YFP⁺ cells from *Csf1r^{flCre} Rosa26^{YFP}* embryos (E8.25, *n* = 7; E8.5, *n* = 4; E9.25–E9.5, *n* = 16; E10.25, *n* = 9; E10.5, *n* = 5; E11.5, *n* = 8; E12.5, *n* = 5). **d**, Number

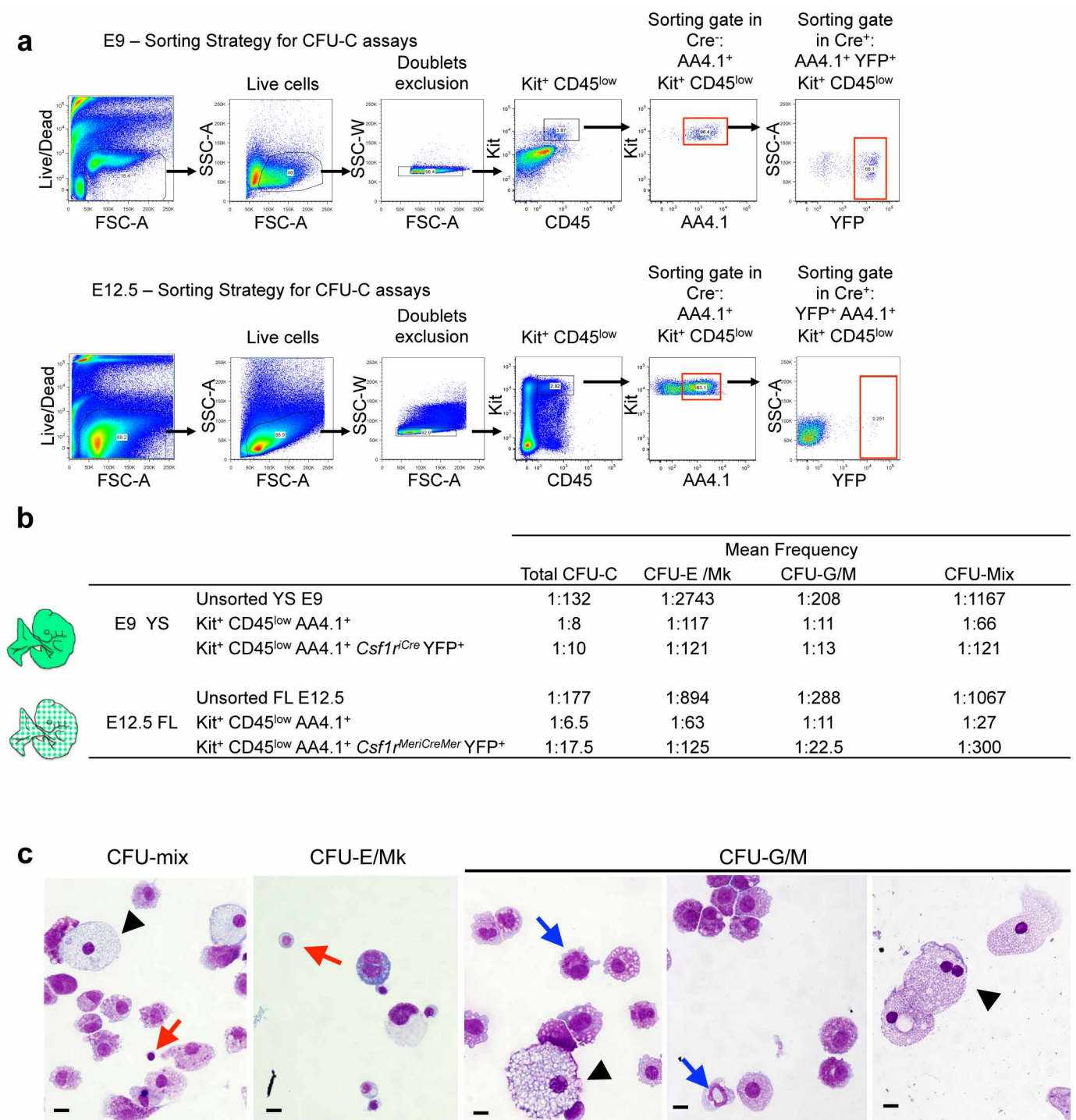
of YFP⁺ Kit⁺ CD45^{low} cells per organ/region and developmental time points (mean ± s.e.m.) in *Csf1r^{flCre} Rosa26^{YFP}* embryos (upper panel). Number of YFP⁺ AA4.1⁺ Kit⁺ CD45^{low} cells per embryonic region and developmental time points (mean ± s.e.m.) in *Csf1r^{flCre} Rosa26^{YFP}* embryos (lower panel). **e**, AA4.1 and Kit expression on YFP⁺ cells from *Csf1r^{flCre} Rosa26^{YFP}* embryos (upper panel) and from *Csf1r^{MerCreMer} Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5 (lower panel).



Extended Data Figure 2 | Fate-mapping analysis of *Csf1r*-expressing cells.

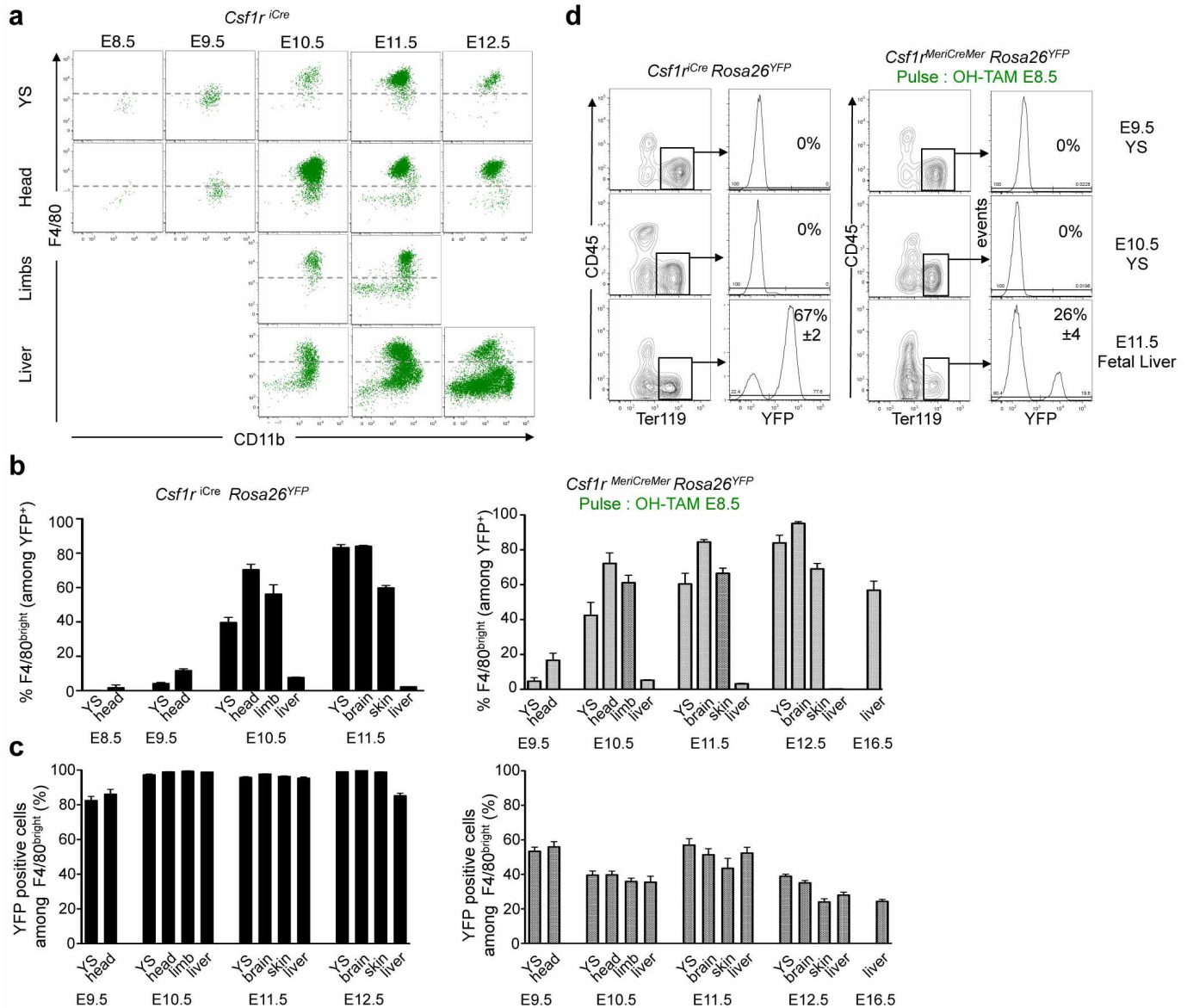
a, Experimental design for fate-mapping analysis of *Csf1r*-expressing cells. Arrows indicate analysed time points. **b**, YFP expression on live cells from *Csf1r^{MeriCreMer} Rosa26^{YFP}* embryos pulsed at E6.5 with OH-TAM and analysed at E10.5 ($n = 2$) and E12.5 ($n = 4$). **c**, Percentage of YFP⁺ cells among Kit⁺ CD45^{low} cells (YFP labelling efficiency) per organ/region (mean \pm s.e.m.). Upper panel, *Csf1r^{Cre} Rosa26^{YFP}* embryos (E8.25, $n = 7$; E8.5, $n = 4$; E9.5,

$n = 16$; E10.25, $n = 9$; E10.5, $n = 5$; E11.5, $n = 8$; E12.5, $n = 5$); lower panel, *Csf1r^{MeriCreMer} Rosa26^{YFP}* embryos pulsed at E8.5 (E9.5, $n = 3$; E10.25, $n = 3$; E10.5, $n = 4$; E11.5, $n = 4$; E12.5, $n = 9$). **d**, Percentage of YFP⁺ cells among AA4.1⁺ Kit⁺ CD45^{low} cells (YFP labelling efficiency) per embryonic organ/region and developmental time points (mean \pm s.e.m.). Upper panel, *Csf1r^{Cre} Rosa26^{YFP}* embryos; lower panel, *Csf1r^{MeriCreMer} Rosa26^{YFP}* embryos pulsed at E8.5.



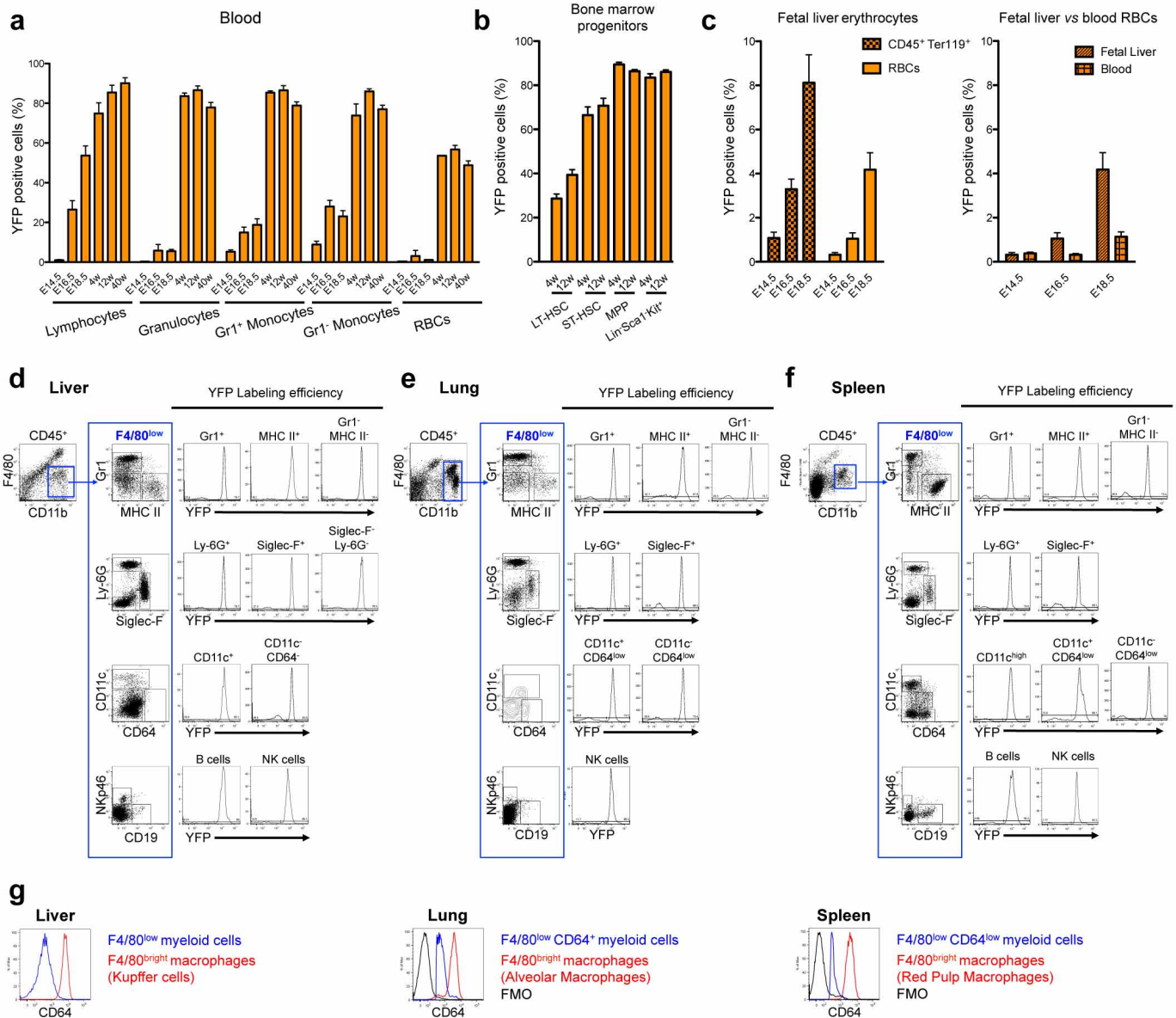
Extended Data Figure 3 | *Csf1*⁺ progenitors have erythro-myeloid potential *ex vivo*. **a**, Sorting strategy for CFU-C (colony forming unit-culture) assays for E9 *Csf1*^{Cre}*Rosa26*^{YFP} yolk sac (upper panel) and E12.5 fetal liver from *Csf1*^{MerCreMer}*Rosa26*^{YFP} embryos pulsed with OH-TAM at E8.5 (lower panel). Dead cells were excluded based on Hoechst 33258 incorporation and, after doublet exclusion, cells were gated based on CD45 and Kit expression. AA4.1⁺Kit⁺CD45^{lo} and YFP⁺AA4.1⁺Kit⁺CD45^{lo} cells were isolated from Cre⁻ and Cre⁺ embryos respectively. **b**, Mean CFU-C frequency from three independent experiments each of E9 *Csf1*^{Cre}*Rosa26*^{YFP} YS and E12.5 fetal liver

from *Csf1*^{MerCreMer}*Rosa26*^{YFP} embryos pulsed with OH-TAM at E8.5. CFU-erythroid and/or megakaryocyte (E/Mk); CFU-granulocyte and/or monocyte/macrophage (G/M); CFU-mix, at least three of the following: G, E, M and Mk. **c**, Morphological validation of colony types obtained from E9 yolk sac *Csf1*^{Cre} YFP⁺ AA4.1⁺Kit⁺CD45^{lo} CFU-C assays. Representative images from May-Grünwald-Giemsa stained cytospin preparations of mixed, E/Mk and G/M colonies. Black arrowhead, macrophages; granulocyte pathway, blue arrows; erythroid and megakaryocyte pathway, red arrows. Scale bar, 10 μm.



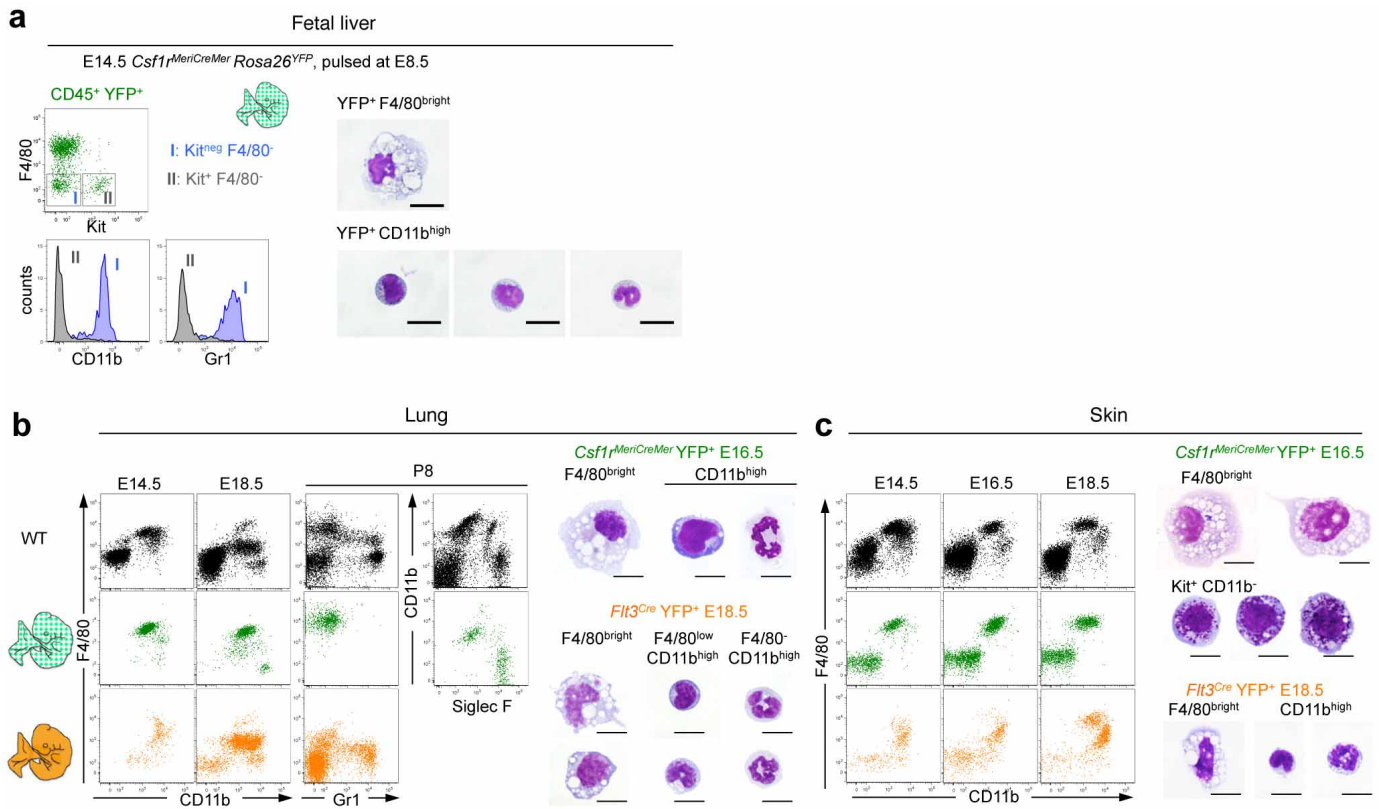
Extended Data Figure 4 | Analysis of *Csf1r* reporter expression in fetal macrophages and red blood cells in *Csf1r^{iCre}Rosa26^{YFP}* embryos. **a**, F4/80 and CD11b expression on YFP⁺ CD45⁺ from yolk sac, head (brain for E11.5), limbs and liver of *Csf1r^{iCre}Rosa26^{YFP}* embryos (E8.5, *n* = 4; E9.5, *n* = 16; E10.5, *n* = 5; E11.5, *n* = 9; E12.5, *n* = 5). Dashed line represents FMO (fluorescence minus one) control. **b**, Percentage of macrophages (F4/80^{bright}) among YFP⁺ cells, mean ± s.e.m., in *Csf1r^{iCre}Rosa26^{YFP}* embryos (left) and in *Csf1r^{MeriCreMer}Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5 (right). See also Supplementary Table 1. **c**, Percentage of YFP⁺ cells among F4/80^{bright}

cells (YFP labelling efficiency) per embryonic organ/region and developmental time points (mean ± s.e.m.). Left panel, *Csf1r^{iCre}Rosa26^{YFP}* embryos (E8.25, *n* = 7; E8.5, *n* = 5; E9.5, *n* = 15; E10.25, *n* = 9; E10.5, *n* = 5; E11.5, *n* = 9; E12.5, *n* = 5); right panel: *Csf1r^{MeriCreMer}Rosa26^{YFP}* embryos pulsed at E8.5 (E9.5, *n* = 3; E10.25, *n* = 3; E10.5, *n* = 4; E11.25, *n* = 4; E12.5, *n* = 9). **d**, YFP expression in erythrocytes (CD45⁺Ter119⁺) from yolk sac and fetal liver of *Csf1r^{iCre}Rosa26^{YFP}* embryos (left) and *Csf1r^{MeriCreMer}Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5 (right).



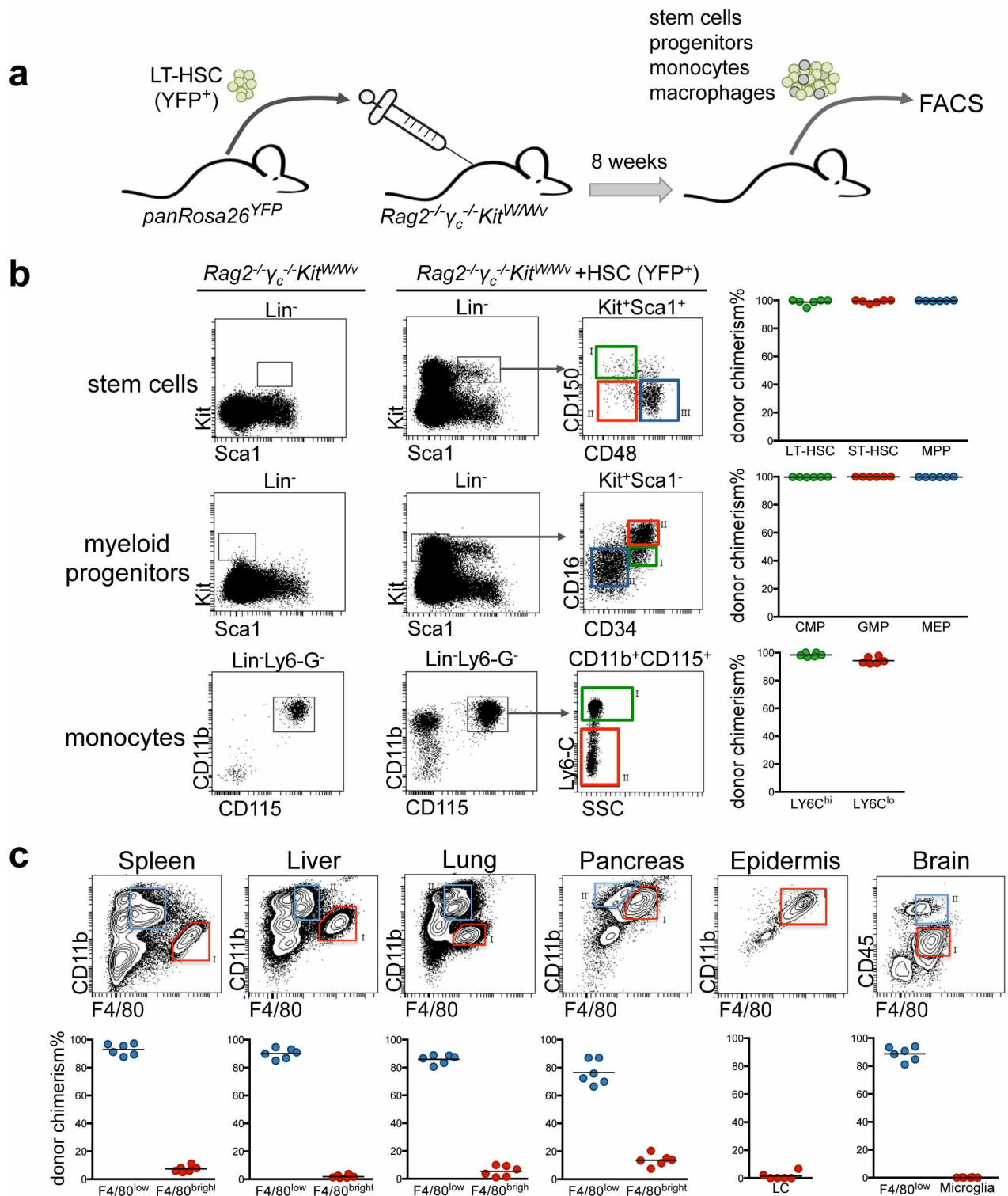
Extended Data Figure 5 | Analysis of *Flt3* reporter expression in blood leucocytes, stem/progenitor cells, fetal red blood cells, and adult liver, lung and spleen in *Flt3*^{Cre}*Rosa26*^{YFP} mice. **a**, YFP labelling efficiency in blood lineages at different embryonic and adult time points (E14.5, *n* = 9; E16.5, *n* = 9; E18.5, *n* = 7; P8, *n* = 7; 4-week-old, *n* = 6; 12-week-old, *n* = 9; 40-week-old, *n* = 7) in *Flt3*^{Cre}*Rosa26*^{YFP} mice are shown. Lymphocytes were gated as CD3⁺/CD19⁺, granulocytes (CD11b⁺Gr1⁺CD115⁻), Gr1⁺ monocytes (CD11b⁺Gr1⁺CD115⁺), Gr1⁻ monocytes (CD11b⁺Gr1⁻CD115⁺) and red blood cells (RBCs, CD45⁺Ter119⁺). **b**, YFP labelling efficiency in bone marrow LT-HSCs, ST-HSCs, MPPs and Lin⁻Sca1⁻Kit⁺ progenitors in 4-week-old (*n* = 3) and 12-week-old (*n* = 6) *Flt3*^{Cre}*Rosa26*^{YFP} mice. **c**, YFP labelling efficiency in fetal liver red blood cell progenitors (CD45⁺Ter119⁺) and red

blood cells (CD45⁻Ter119⁺) in *Flt3*^{Cre}*Rosa26*^{YFP} mice (E14.5, *n* = 5; E16.5, *n* = 5; E18.5, *n* = 7), and comparison of YFP labelling efficiency in fetal liver and blood red blood cells in *Flt3*^{Cre}*Rosa26*^{YFP} mice at E14.5 (*n* = 5), E16.5 (*n* = 5) and E18.5 (*n* = 7). **d**, Expression of Gr1 and MHC II, Ly-6G and Siglec-F, CD11c and CD64, and Nkp46 and CD19 among F4/80^{lo}CD11b^{hi} myeloid cells in the liver. Histograms represent *Flt3*^{Cre} YFP labelling efficiency in the following defined populations: granulocytes (Gr1⁺MHC II⁻ or Ly-6G⁺), eosinophils (Siglec-F⁺), dendritic cells (CD11c⁺), B cells (CD19⁺) and NK cells (Nkp46⁺) (*n* = 3). **e**, Analysis of F4/80^{lo}CD11b^{hi} myeloid cells in the lung as in **b**. **f**, Analysis of F4/80^{lo}CD11b^{hi} myeloid cells in the spleen as in **b**. **g**, Expression of CD64 in F4/80^{bright} macrophages and in F4/80^{lo} myeloid cells in the liver, lung and spleen (FMO, fluorescence minus one).



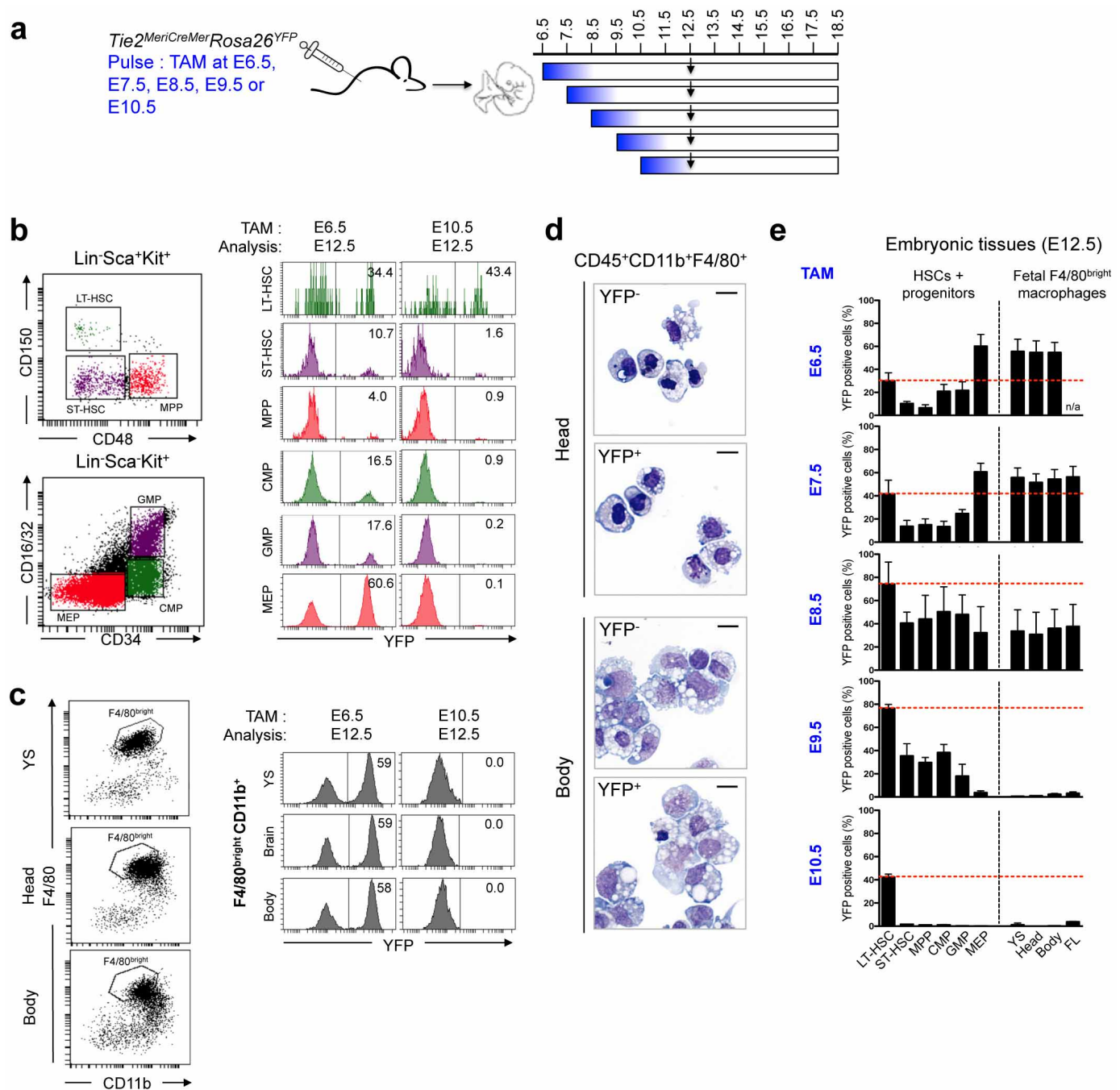
Extended Data Figure 6 | Characterization of fetal F4/80^{lo}CD11b^{hi} myeloid cells in liver, lung and skin. **a**, F4/80, Kit, CD11b and Gr1 expression on YFP⁺CD45⁺ cells in the fetal liver at E14.5 in *Csf1^{MerCreMer}Rosa26^{YFP}* embryos pulsed at E8.5 (left panel). Representative images of May-Grünwald-Giemsa stained cytopsin preparations of fetal liver YFP⁺F4/80^{bright} and YFP⁺CD11b^{hi} cells sorted from E14.5 *Csf1^{MerCreMer}Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5 (right panel). Scale bar, 10 μ m. **b**, **c**, F4/80, CD11b, Gr1 and Siglec F expression on CD45⁺ cells in the embryonic and

post-natal lung (**b**) and skin (**c**) in *Csf1^{MerCreMer}Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5 (green) and *Flt3^{Cre}Rosa26^{YFP}* embryos (orange). Representative images of May-Grünwald-Giemsa stained cytopsin preparations of lung YFP⁺F4/80^{bright} and YFP⁺CD11b^{hi}F4/80^{lo} (**b**) and skin YFP⁺F4/80^{bright} and YFP⁺Kit⁺F4/80⁻CD11b⁻ mast cells (**c**) sorted from E18.5 *Flt3^{Cre}Rosa26^{YFP}* embryos and E16.5 *Csf1^{MerCreMer}Rosa26^{YFP}* embryos pulsed with OH-TAM at E8.5. Scale bar, 10 μ m.



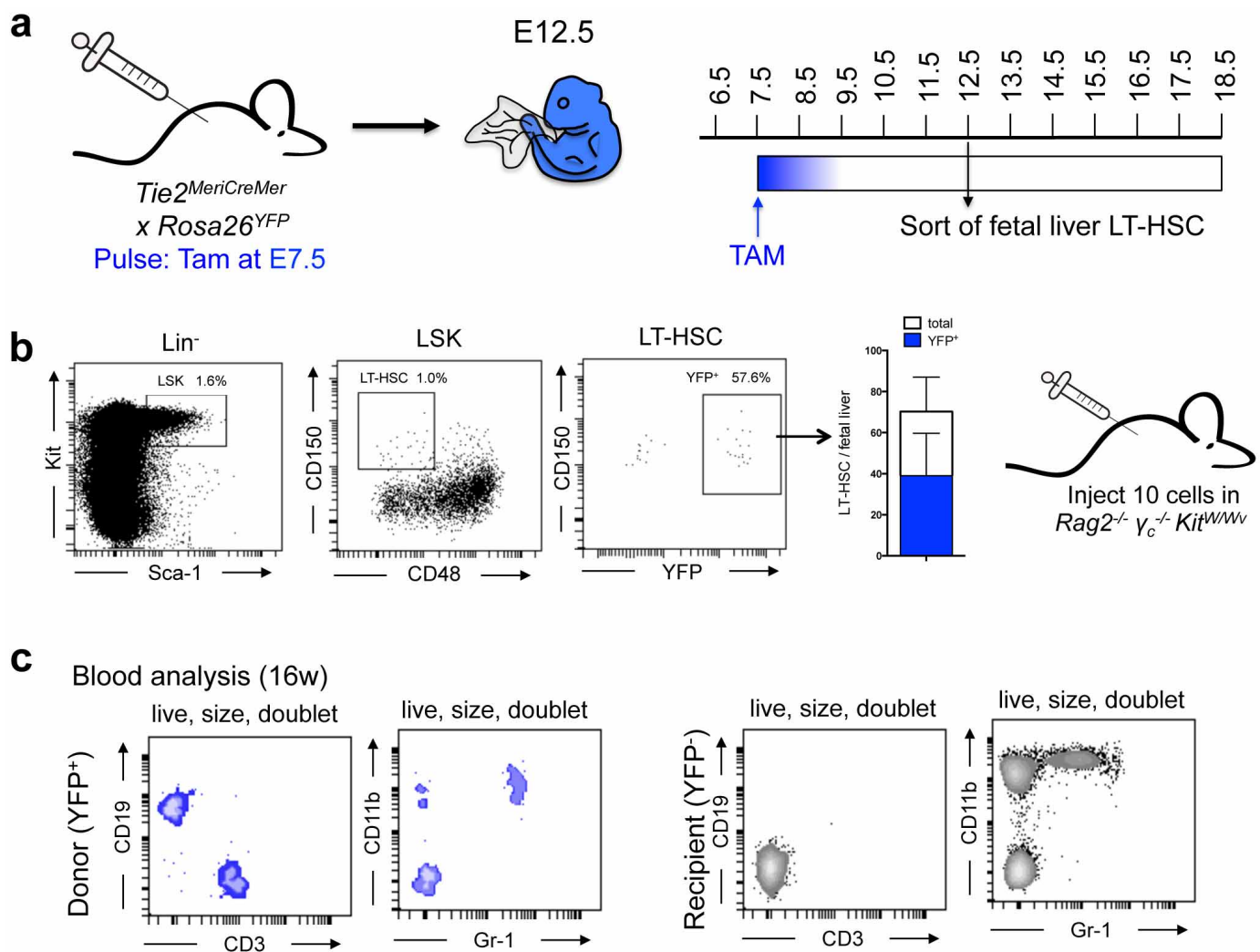
Extended Data Figure 7 | Adult BM transplantation reconstitutes the haematopoietic system but does not replace tissue-resident F4/80^{bright} macrophages. **a**, Schematic representation of transplantation experiments. LT-HSCs isolated from bone marrow of *panRosa26^{YFP}* donor mice were injected into *Rag2^{-/-}γc^{-/-}Kit^{W/Wv}* recipients (approximately 1,000 cells per recipient). Eight weeks after transplantation stem cells, myeloid progenitors, monocytes and macrophages of recipient mice were analysed for donor chimerism. **b**, Long-term or short-term haematopoietic stem cells (LT-HSCs,

ST-HSCs), multipotent progenitors (MPPs), common myeloid progenitors (CMPs), granulocyte-monocyte progenitors (GMPs), megakaryocyte-erythrocyte progenitors (MEPs), and circulating Ly6C^{hi} and Ly6C^{lo} monocytes were isolated from transplanted *Rag2^{-/-}γc^{-/-}Kit^{W/Wv}* mice and analysed for YFP expression. **c**, F4/80^{bright} macrophages and F4/80^{lo} myeloid cells in spleen, liver, lung, pancreas, epidermis and brain were analysed for YFP expression.



Extended Data Figure 8 | Analysis of fetal stem/progenitor cells and fetal macrophages in *Tie2^{MerCreMer}Rosa26^{YFP}* embryos pulse-labelled from E6.5 to E10.5. a, Experimental design for fate-mapping analysis of *Tie2^{MerCreMer}Rosa26^{YFP}* embryos pulse-labelled at E6.5, or E7.5, or E8.5, or E9.5 or E10.5. b, c, Representative flow cytometry of fetal liver stem/progenitor cells (b) and of fetal macrophages (c) in the yolk sac, head region, and embryo body at E12.5, injected at E6.5 or at E10.5. d, Representative images of

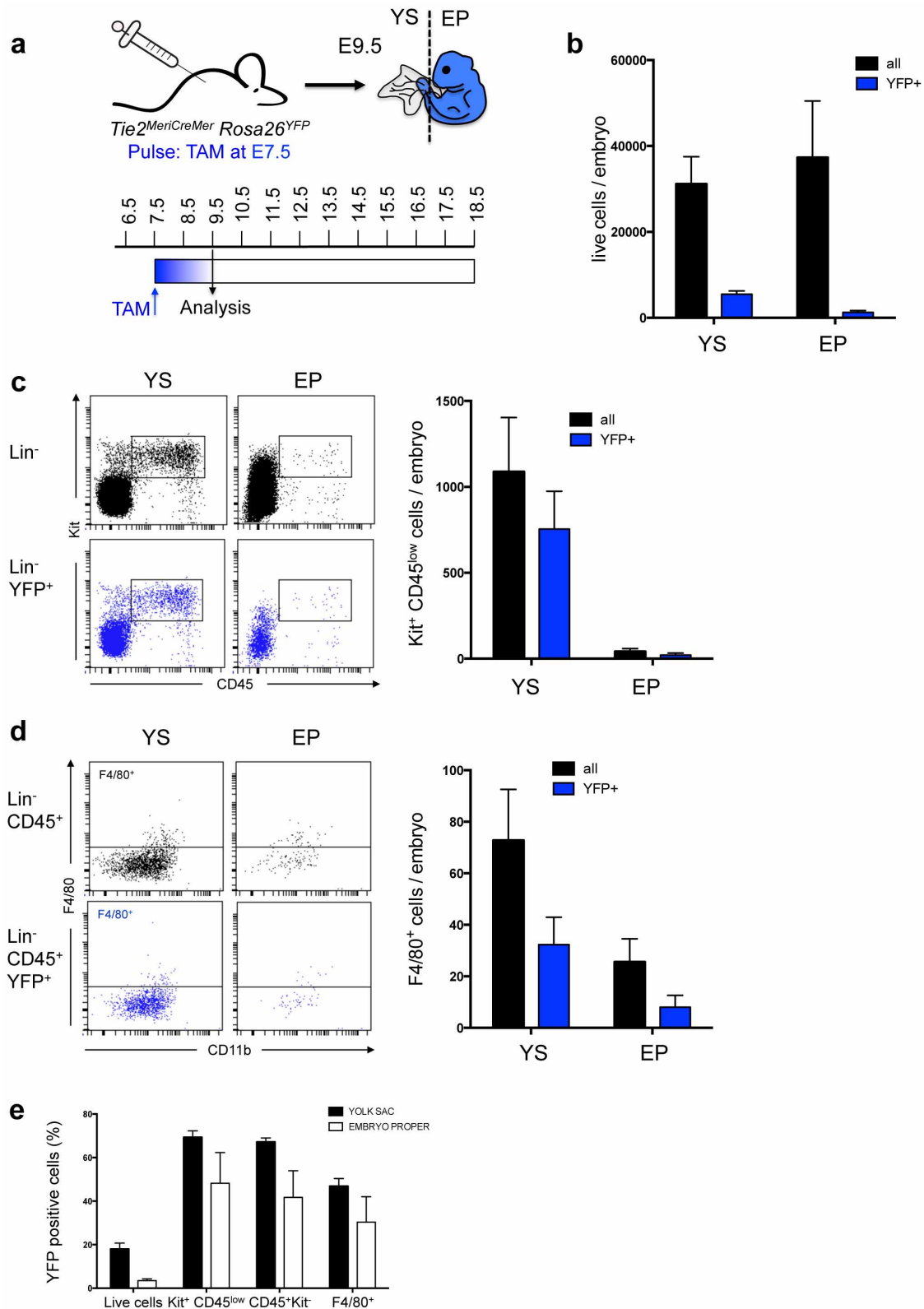
May-Grünwald-Giemsa stained cytospin preparations of sorted YFP⁺ and YFP⁻ CD45⁺ F4/80^{bright} macrophages of the embryo proper or the head region of E13.5 *Tie2^{MerCreMer}Rosa26^{YFP}* embryos pulsed at E7.5. Scale bar, 10 μ m. e, Quantification of the percentage of YFP⁺ stem/progenitor cells in the fetal liver and macrophages in yolk sac, brain (head) and embryo body at E12.5. Embryos were labelled at E6.5 ($n = 5$), or E7.5 ($n = 7$), or E8.5 ($n = 4$), or E9.5 ($n = 5$) or E10.5 ($n = 7$) and analysed at E12.5 (mean \pm s.d.).



Extended Data Figure 9 | Transplantation of YFP⁺ fetal liver LT-HSCs from *Tie2^{MeriCreMer} Rosa26^{YFP}* into *Rag2^{-/-} γ_c^{-/-} Kit^{W/Wv}* mice.

a, Experimental design for pulse-labelling and LT-HSCs sort from *Tie2^{MeriCreMer} Rosa26^{YFP}* embryos pulsed at E7.5. **b**, Fetal livers of E12.5 embryos pulsed at E7.5 were collected, YFP⁺ LSK CD150⁺ CD48⁻ LT-HSCs

were sorted by flow cytometry and 10 LT-HSCs were injected into *Rag2^{-/-} γ_c^{-/-} Kit^{W/Wv}* recipients. **c**, Blood analysis of recipients 16 weeks after LT-HSC transplantation. Donor-derived (YFP⁺) and recipient-derived (YFP⁻) blood cells were analysed for expression of CD19, CD3, CD11b, and Gr-1. One representative example is shown.



Extended Data Figure 10 | Analysis of E9.5 YS progenitor cells in *Tie2^{MerCreMer} Rosa26^{YFP}* embryos pulse-labelled at E7.5. **a**, Experimental design for fate-mapping analysis of *Tie2*-expressing cells in *Tie2^{MerCreMer} Rosa26^{YFP}* embryos. Embryonic cells were pulse-labelled by tamoxifen (TAM) administration into pregnant *Tie2^{MerCreMer}* mice at E7.5. Yolk sac (YS) and embryo proper (EP) of E9.5 *Tie2^{MerCreMer} Rosa26^{YFP}* embryos were analysed by flow cytometry. **b**, Quantification of total living cells and YFP⁺ living cells in yolk sac and embryo proper of analysed embryos (mean \pm s.d., $n = 5$). **c**, Flow cytometry analysis of Kit⁺ CD45^{low} cells among

total living cells (black) or living YFP⁺ cells (blue) in yolk sac and embryo proper of a representative E9.5 *Tie2^{MerCreMer} Rosa26^{YFP}* embryo (left) and quantification of all analysed embryos (right; mean \pm s.d., $n = 5$). **d**, Analysis of F4/80⁺ fetal macrophages among CD45⁺ cells (black) or YFP⁺ CD45⁺ cells (blue) in yolk sac and embryo proper (mean \pm s.d., $n = 5$) and quantification of all analysed embryos (right; mean \pm s.d., $n = 5$). **e**, Percentage of YFP⁺ cells (YFP labelling efficiency) among live cells, Kit⁺ CD45^{low}, CD45⁺ Kit⁻ cells and F4/80⁺ cells from the yolk sac and embryo proper of E9.5 *Tie2^{MerCreMer} Rosa26^{YFP}* embryos pulsed at E7.5.

Role of *TP53* mutations in the origin and evolution of therapy-related acute myeloid leukaemia

Terrence N. Wong^{1*}, Giridharan Ramsingh^{2*}, Andrew L. Young^{3*}, Christopher A. Miller⁴, Waseem Touma¹, John S. Welch^{1,5}, Tamara L. Lamprecht¹, Dong Shen⁶, Jasreet Hundal⁴, Robert S. Fulton⁴, Sharon Heath¹, Jack D. Baty⁷, Jeffery M. Klcó⁸, Li Ding^{1,5}, Elaine R. Mardis^{4,5,9}, Peter Westervelt^{1,5}, John F. DiPersio^{1,5}, Matthew J. Walter^{1,5}, Timothy A. Graubert^{1,5}, Timothy J. Ley^{1,5}, Todd E. Druley³, Daniel C. Link^{1,5} & Richard K. Wilson^{4,5,9}

Therapy-related acute myeloid leukaemia (t-AML) and therapy-related myelodysplastic syndrome (t-MDS) are well-recognized complications of cytotoxic chemotherapy and/or radiotherapy¹. There are several features that distinguish t-AML from *de novo* AML, including a higher incidence of *TP53* mutations^{2,3}, abnormalities of chromosomes 5 or 7, complex cytogenetics and a reduced response to chemotherapy⁴. However, it is not clear how prior exposure to cytotoxic therapy influences leukaemogenesis. In particular, the mechanism by which *TP53* mutations are selectively enriched in t-AML/t-MDS is unknown. Here, by sequencing the genomes of 22 patients with t-AML, we show that the total number of somatic single-nucleotide variants and the percentage of chemotherapy-related transversions are similar in t-AML and *de novo* AML, indicating that previous chemotherapy does not induce genome-wide DNA damage. We identified four cases of t-AML/t-MDS in which the exact *TP53* mutation found at diagnosis was also present at low frequencies (0.003–0.7%) in mobilized blood leukocytes or bone marrow 3–6 years before the development of t-AML/t-MDS, including two cases in which the relevant *TP53* mutation was detected before any chemotherapy. Moreover, functional *TP53* mutations were identified in small populations of peripheral blood cells of healthy chemotherapy-naïve elderly individuals. Finally, in mouse bone marrow chimaeras containing both wild-type and *Tp53*^{+/-} haematopoietic stem/progenitor cells (HSPCs), the *Tp53*^{+/-} HSPCs preferentially expanded after exposure to chemotherapy. These data suggest that cytotoxic therapy does not directly induce *TP53* mutations. Rather, they support a model in which rare HSPCs carrying age-related *TP53* mutations are resistant to chemotherapy and expand preferentially after treatment. The early acquisition of *TP53* mutations in the founding HSPC clone probably contributes to the frequent cytogenetic abnormalities and poor responses to chemotherapy that are typical of patients with t-AML/t-MDS.

t-AML and t-MDS are clonal haematopoietic disorders that typically develop 1–5 years after exposure to chemotherapy or radiotherapy¹. To understand better how prior cytotoxic therapy contributes to the high incidence of *TP53* mutations and karyotypic abnormalities in t-AML/t-MDS, we sequenced the genomes of 22 cases of t-AML, including one case that has been previously reported⁵. These data were compared to whole-genome sequence data previously reported for *de novo* AML⁶ and secondary AML (s-AML) arising from MDS for which patients did not receive chemotherapy except hydroxyurea^{7,8}. Of the sequenced t-AML cases, 23% had rearrangements of MLL (also known as *KMT2A*), 23% had complex cytogenetics, and 36% had normal cytogenetics (Extended Data Table 1 and Supplementary Table 1).

We predicted that DNA damage induced during exposure to cytotoxic therapy would manifest itself in t-AML genomes with an increased

mutation burden. However, the total number of validated somatic single-nucleotide variants (SNVs) and genic (tier 1) somatic SNVs identified was similar to that for *de novo* AML and s-AML (Fig. 1a, b). Likewise, the number of small insertions or deletions (indels) in genic regions was similar in t-AML, *de novo* AML and s-AML (Fig. 1c). A previous study

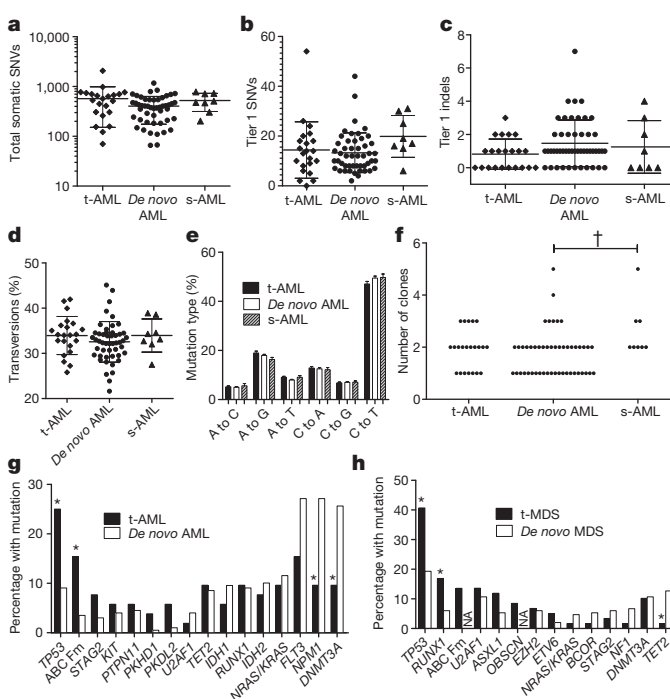


Figure 1 | The mutational burden in t-AML is similar to *de novo* AML.

a, Total number of validated tier 1–3 somatic SNVs in t-AML ($n = 22$), *de novo* AML ($n = 49$) and s-AML ($n = 8$). The mean ages of the t-AML, *de novo* AML and s-AML cohorts were 55.7, 51 and 54.6 years, respectively. **b**, Number of validated tier 1 somatic SNVs. **c**, Number of validated tier 1 small indels. **d**, Percentage of tier 1–3 somatic SNVs that are transversions. **e**, Mutational spectrum for all validated tier 1–3 somatic SNVs. **f**, Number of distinct clones per sample inferred from the identification of discrete clusters of mutations with distinct variant allele frequencies. **g**, Percentage of cases of t-AML ($n = 52$) or *de novo* AML ($n = 199$) harbouring non-synonymous mutations of the indicated gene. **h**, Percentage of cases of t-MDS ($n = 59$) or *de novo* MDS ($n = 150$) harbouring non-synonymous mutations of the indicated gene. ABC Fm, ABC family genes; NA, not available. † $P < 0.05$ by one-way analysis of variance (ANOVA). * $P < 0.05$ by Fisher's exact test. Data represent the mean \pm standard deviation (s.d.).

¹Department of Medicine, Division of Oncology, Washington University, St Louis, Missouri 63110, USA. ²Department of Medicine, Jane Anne Nohl Division of Hematology, University of Southern California, Los Angeles, California 90089, USA. ³Department of Pediatrics, Division of Hematology/Oncology, Washington University, St Louis, Missouri 63110, USA. ⁴The Genome Institute, Washington University, St Louis, Missouri 63110, USA. ⁵Siteman Cancer Center, Washington University, St Louis, Missouri 63110, USA. ⁶AstraZeneca, Gaithersburg, Maryland 20878, USA. ⁷Division of Biostatistics, Washington University, St Louis, Missouri 63110, USA. ⁸Department of Pathology and Immunology, Washington University, St Louis, Missouri 63110, USA. ⁹Department of Genetics, Washington University, St Louis, Missouri 63110, USA.

*These authors contributed equally to this work.

showed that transversions are specifically enriched in relapsed AML after chemotherapy⁹. However, the percentage of transversions, and in fact of all six classes of SNVs, was similar in all three cohorts (Fig. 1d, e). Structural variants and somatic copy number alterations were uncommon in these t-AML cases (Supplementary Table 2 and Extended Data Fig. 1a). Moreover, the number of identifiable subclones in t-AML was similar to that observed in *de novo* AML (Fig. 1f and Extended Data Fig. 1b). Collectively, these data show that the mutation burden of t-AML genomes is similar to that of *de novo* AML genomes.

We next asked whether the pattern of genes frequently mutated in t-AML/t-MDS is distinct from that observed in *de novo* AML/MDS. Whole-genome sequencing identified an average of 10.2 ± 7.1 missense, nonsense, in-frame indel or frameshift mutations per t-AML genome (Supplementary Table 3). To define better the frequency of specific mutations in t-AML/t-MDS, we sequenced a panel of 149 AML/MDS-related genes in an additional 89 patients with t-AML or t-MDS (Supplementary Table 4). We combined the whole-genome sequence data with the extension series to report on 52 cases of t-AML and 59 cases of t-MDS. Abnormalities of chromosome 5 or 7 or complex cytogenetics were present in 55.0% of cases (Extended Data Table 2 and Supplementary Table 1). The t-AML/t-MDS data were compared to 199 previously reported *de novo* AML genomes or exomes⁶, or 150 previously reported cases of *de novo* MDS in which extensive candidate gene sequencing was performed⁸. As reported previously, *TP53* mutations are significantly enriched in t-AML/t-MDS compared with *de novo* AML/MDS (Fig. 1g, h and Supplementary Table 5). Interestingly, mutations of ABC transporter genes, a subset of which have been implicated in chemotherapy resistance, are also enriched in t-AML versus *de novo* AML. On the other hand, several well-defined driver gene mutations (that is, *DNMT3A* and *NPM1*) were significantly less common in t-AML. Thus, although the total mutation burden is similar, a distinct subset of mutated genes is present in t-AML/t-MDS.

TP53 is the most commonly mutated gene in t-AML/t-MDS, with 33.3% of patients affected in our series (Fig. 1g, h); the vast majority of these mutations have previously been identified as pathogenic¹⁰. Multivariate analysis revealed that *TP53* mutations were associated with poor risk cytogenetics and a worse prognosis (Supplementary Tables 6, 7 and Extended Data Fig. 2), both hallmarks of t-AML/t-MDS. These observations suggest a central role for *TP53* mutations in the pathogenesis of many cases of t-AML/t-MDS. However, the mechanism by which *TP53*

mutations are selectively enriched in t-AML/t-MDS is unclear. The mutation burden in the genomic region containing *TP53* (including silent tier 1, and any tier 2 or tier 3 mutations) is similar between t-AML and *de novo* AML (Extended Data Fig. 1c). Thus, it is not likely that chemotherapy directly induces *TP53* mutations. We recently reported that individual HSPCs accumulate somatic mutations as a function of age, such that by age 50, there are on average five coding gene mutations per HSPC¹¹. On the basis of these data and on current estimates that there are approximately 10,000 haematopoietic stem cells (HSCs) in humans¹², we predict that 44% of healthy individuals at 50 years of age may have at least one HSPC that carries a randomly generated, functional *TP53* mutation (see Methods). *TP53* has a central role in regulating cellular responses to genotoxic stress^{13–17}, and loss of *TP53* provides a selective advantage for neoplastic growth¹⁸. Together, these observations suggest a model in which rare HSPCs carrying age-related *TP53* mutations are resistant to chemotherapy and expand preferentially after treatment (Extended Data Fig. 3).

This model suggests the following testable predictions: (1) in patients with t-AML containing clonal *TP53* mutations, HSPCs harbouring the specific *TP53* mutation will be present long before the development of overt t-AML; (2) somatic *TP53* mutations will be present in the HSPCs of some healthy individuals never exposed to cytotoxic therapy; and (3) HSPCs harbouring *TP53* mutations will expand under the selective pressure of chemotherapy.

To test the first prediction, we identified seven cases of t-AML/t-MDS with specific *TP53* mutations for which we had leukapheresis or bone marrow specimens banked 3–8 years before the development of t-AML/t-MDS (Extended Data Table 3). Of note, in all the cases, the *TP53* mutation was clonal in the t-AML/t-MDS diagnostic sample. Current next-generation sequencing technology is limited in the detection of rare variant alleles owing to an intrinsic sequencing error rate of $\sim 0.1\%$ (ref. 19). To overcome this limitation, we introduced random barcodes during production of the sequencing libraries, such that sequence 'read families' containing unique barcodes are generated (Extended Data Fig. 4a). Using tumour DNA with a known *TP53* mutation, we show that this assay can detect a variant allele with a frequency of 0.009% (Extended Data Fig. 4b, c).

The specific *TP53* mutation present in the diagnostic t-AML/t-MDS sample was identified in previously banked specimens in four out of the seven cases tested (see Supplementary Notes for case presentations). In

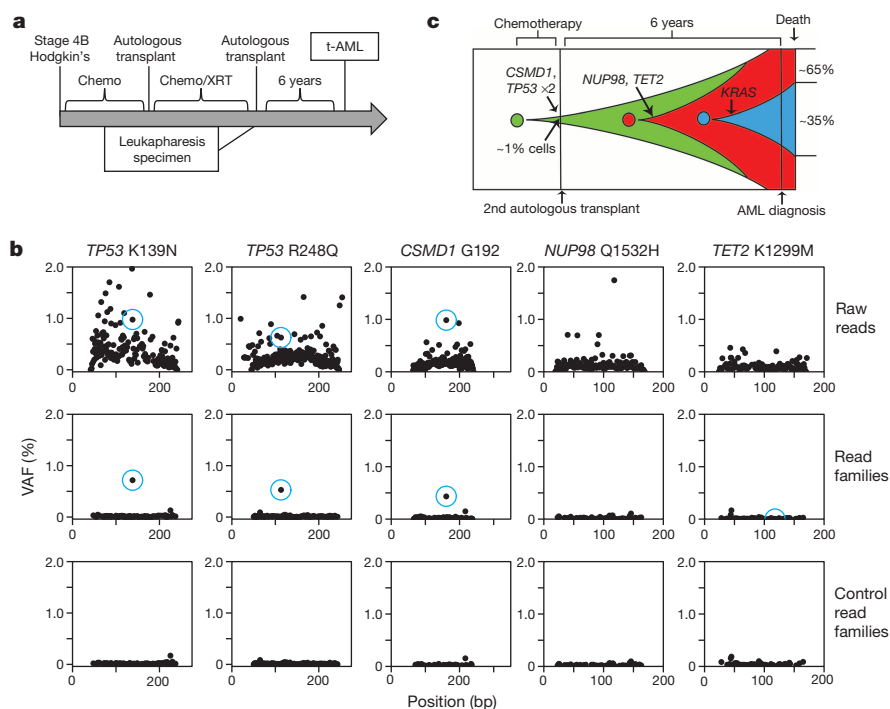


Figure 2 | Biallelic *TP53* mutations are early mutational events in the AML cells of UPN 530447. **a**, Clinical course of case 530447. Chemo, chemotherapy; XRT, radiation therapy. **b**, Unique adaptor sequencing of a leukapheresis sample obtained 6 years before the diagnosis of t-AML for each of the five clonal somatic SNVs identified in the diagnostic t-AML sample. Genomic DNA from a patient lacking these variants served as a control. Blue circles indicate the position of the variant SNV. **c**, Proposed model of clonal evolution to t-AML in this case.

the other three cases, we were unable to detect the diagnostic *TP53* mutation in the previously banked blood or bone marrow sample; it is not clear whether these mutations were present but below our limit of detection or were truly absent. Patient 530447 developed t-AML after an autologous stem cell transplant for refractory Hodgkin's lymphoma (Fig. 2a). The diagnostic t-AML sample carried biallelic mutations of *TP53*, missense mutations of *TET2* and *NUP98*, a silent mutation of *CSMD1*, and a subclonal *KRAS* mutation. Analysis of a leukapheresis sample obtained 6 years before the development of t-AML revealed that both *TP53* mutant alleles were present with a variant allele fraction (VAF) of approximately 0.5% (Fig. 2b). The *CSMD1* mutation was also present at the same VAF and is probably a passenger mutation. However, two potential driver mutations (*TET2* and *NUP98*) were not detectable in the previously banked sample. Thus, these data show that, in this patient, the biallelic *TP53* mutations preceded the development of t-AML by at least 6 years and antedated the development of the *TET2* and *NUP98* mutations (Fig. 2c). In a second case (unique patient number (UPN) 341666), a heterozygous *TP53* R196* mutation was identified in mobilized peripheral blood leukocytes 3 years before the development of t-MDS at a frequency of 0.1%, preceding the acquisition of a *RUNX1* mutation (Extended Data Fig. 5).

In two of the four cases, the previously banked sample was obtained before the initiation of chemotherapy. Patient 967645 developed t-AML 5 years after the diagnosis of marginal zone lymphoma (Fig. 3a). The diagnostic t-AML sample contained a homozygous *TP53* Y220C mutation. Using a droplet digital polymerase chain reaction (ddPCR) assay, we identified the same *TP53* Y220C mutation in a bone marrow sample obtained before any chemotherapy at a frequency of 0.0027% (average of two independent experiments) (Fig. 3b). We next asked whether other mutations in the diagnostic t-AML sample were also present in this previously banked sample (Supplementary Table 8). We focused on the G155S mutation in *SNAP25*; this mutation is probably non-pathogenic as *SNAP25* is not expressed in AML samples⁶. Indeed, we identified the *SNAP25* G155S mutation in the previously banked bone marrow sample with a similar VAF (0.0029%) as that for *TP53* Y220C (Fig. 3c). Of note, deletion (del)(5q) and del(7q) were subclonal at diagnosis (present in 54% and 38% of metaphases, respectively) (Supplementary Table 1). Collectively, these data provide evidence that an HSPC harbouring a *TP53* Y220C mutation preferentially expanded after chemotherapy with the subsequent acquisition of del(5q) and then del(7q) (Fig. 3d). Of note, we found two other cases of t-AML/t-MDS with clonal *TP53* mutations but subclonal del(5q), del(5) and/or del(7) (UPNs 756582 and 837334, Supplementary Table 1). Together, these data suggest that *TP53* mutations precede the development of these characteristic cytogenetic abnormalities of t-AML/t-MDS.

In a second case, patient 895681 developed t-MDS 5.5 years after the initiation of chemotherapy for non-Hodgkin's lymphoma (Fig. 3e). The diagnostic t-MDS sample contained a clonal *TP53* H179L mutation. Using ddPCR, we identified *TP53* H179L at a VAF of 0.05% in a bone marrow sample taken before the initiation of cytotoxic therapy (Fig. 3f). Thus, as with patient 967645, an HSPC carrying a functional *TP53* mutation was present before cytotoxic therapy exposure, later giving rise to the malignant t-AML/t-MDS clone (Fig. 3g).

To determine whether HSPCs harbouring *TP53* mutations are present in healthy individuals, we analysed peripheral blood leukocytes from 20 elderly (68–89 years old) cancer-free donors who had not received prior cytotoxic therapy. We limited our sequencing to exons 4–8 of *TP53* since the majority of pathogenic mutations in *TP53* are located in these exons. Using our unique adaptor sequencing assay, we identified *TP53* mutations in 9 of 19 evaluable cases, with VAFs ranging from 0.01% to 0.37% (Extended Data Table 4). Of note, since we did not sequence the entire coding region of *TP53*, it is likely that our study underestimates the true frequency of healthy elderly individuals harbouring HSPCs with *TP53* mutations. ddPCR confirmed the presence of the *TP53* mutation in all three cases that were tested (Extended Data Fig. 6). Interestingly, the majority of the *TP53* mutations identified are known pathogenic

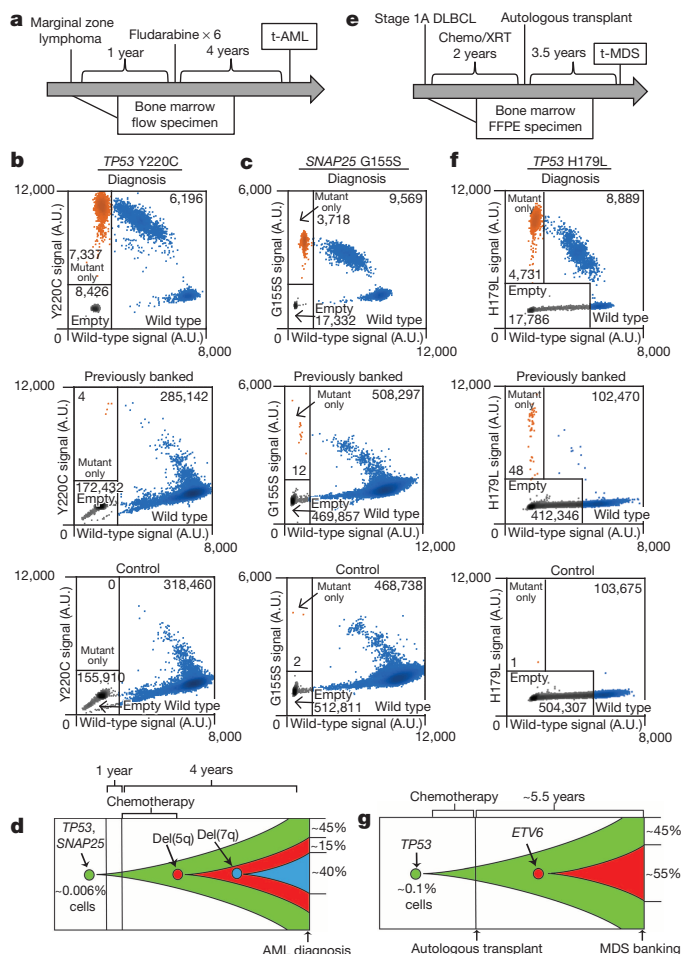


Figure 3 | HSPC clones harbouring somatic *TP53* mutations are detected in patients before cytotoxic therapy exposure. **a**, Clinical course of case 967645. **b**, Dot plots of ddPCR of a diagnostic t-AML sample from case 967645, a bone marrow sample from this patient obtained 5 years before the development of t-AML (before any cytotoxic therapy), or a control sample from a patient lacking a mutation in *TP53*. Droplets containing only the *TP53* Y220C allele are highlighted in orange, droplets containing wild-type *TP53* (with or without *TP53* Y220C) are highlighted in blue; empty droplets are grey. The number of droplets in each gate is indicated. Data are representative of two independent experiments. **c**, Dot plots of ddPCR data for *SNAP25* G155S using the same genomic DNA as in **b**. **d**, Proposed model of clonal evolution to t-AML in case 967645. **e**, Clinical course of case 895681. Chemo, chemotherapy; DLBCL, diffuse large B-cell lymphoma; FFPE, formalin-fixed paraffin-embedded; XRT, radiotherapy. **f**, Dot plots of ddPCR data of the diagnostic t-MDS sample from case 895681, a bone marrow FFPE sample from this patient obtained 5.5 years before the development of t-MDS (before any cytotoxic therapy), or a control FFPE sample obtained from a patient lacking a mutation in *TP53*. The labelling scheme is the same as in **b**. **g**, Proposed model of clonal evolution to t-MDS in case 895681; the diagnostic t-MDS sample contained a subclonal *ETV6* mutation.

mutations previously implicated in cancer. These data suggest that functional *TP53* mutations may confer (even in the absence of cytotoxic therapy) a subtle competitive advantage that results in modest HSPC expansion over time.

To test directly the hypothesis that functional *TP53* mutations confer a competitive advantage after chemotherapy, we generated mixed bone marrow chimaeras containing both wild-type and *TP53*^{+/-} cells (Fig. 4a). In mice treated with vehicle control, we observed a non-significant trend towards an increased *TP53*^{+/-} donor contribution to haematopoiesis (Fig. 4b–e). Whether longer follow-up would confirm a subtle competitive advantage, as suggested by the expansion of *TP53* mutant HSPC clones in elderly healthy individuals, will require additional study. Regardless,

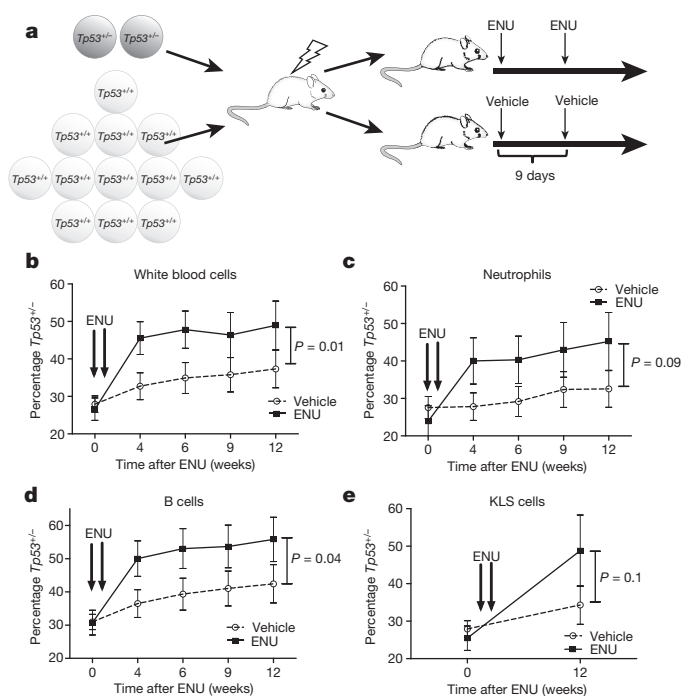


Figure 4 | Heterozygous loss of *TP53* confers a clonal advantage to HSCs after exposure to ENU. **a**, Experimental schema. Bone marrow chimaeras were generated by transplanting a 7 to 1 ratio of wild-type to *TP53*^{+/-} bone marrow into irradiated syngeneic recipients. After hematopoietic reconstitution (5 weeks), mice were treated with ENU or vehicle control as indicated. **b–d**, Shown is the percentage of total leukocytes (**b**), Gr-1⁺ neutrophils (**c**) or B220⁺ B cells (**d**) that were derived from *TP53*^{+/-} cells. **e**, Percentage of Kit⁺ lineage⁻ Sca⁺ (KSL) cells in the bone marrow 12 weeks after ENU exposure that were derived from *TP53*^{+/-} cells. Data represent the mean \pm standard error of the mean (s.e.m.) of 11 and 14 mice in the ENU and vehicle cohorts, respectively. Peripheral chimaerism was analysed using two-way ANOVA and KLS chimaerism was analysed using an analysis of covariance (ANCOVA).

upon treatment with *N*-ethyl-*N*-nitrosourea (ENU), *TP53*^{+/-} HSPCs show a competitive advantage. Importantly, a previous study similarly showed that *TP53*^{+/-} HSCs also have a competitive advantage after irradiation, which appeared to be due, at least in part, to reduced irradiation-induced senescence in *TP53*^{+/-} HSCs^{20,21}.

There is increasing evidence that cancers undergo clonal evolution under the selective pressure of chemotherapy²². For example, the clonal architecture of *de novo* AML is dynamic, with certain (often minor) sub-clones becoming dominant at relapse after chemotherapy⁹. We show that HSPCs that acquire heterozygous *TP53* mutations as a function of normal ageing are also subject to Darwinian selection upon exposure to cytotoxic therapy, ultimately resulting in the expansion of HSPCs with these mutations. The high frequency (nearly 50%) of elderly individuals with detectable heterozygous *TP53* mutations in their circulating leukocytes far exceeds the prevalence of AML or MDS in this age group. Clearly, additional mutations, including mutation of the second *TP53* allele, are needed for transformation to AML or MDS. Consistent with this observation, only a minority of patients with Li–Fraumeni syndrome, most of whom harbour germline heterozygous *TP53* mutations, develop AML or MDS²³. This model provides a potential mechanism for the high incidence of *TP53* mutations in t-AML/t-MDS²⁴. The *TP53* mutation in the founding clone probably contributes to the frequent cytogenetic abnormalities and poor response to chemotherapy that are typical of t-AML/t-MDS. For t-AML/t-MDS cases that do not harbour *TP53* mutations, it will be important to determine whether different age-related mutations also confer a competitive advantage to HSPCs that are exposed to cytotoxic therapy, and to define the nature of these mutations.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 December 2013; accepted 13 October 2014.

Published online 8 December 2014.

- Godley, L. A. & Larson, R. A. Therapy-related myeloid leukemia. *Semin. Oncol.* **35**, 418–429 (2008).
- Christiansen, D. H., Andersen, M. K. & Pedersen-Bjergaard, J. Mutations with loss of heterozygosity of p53 are common in therapy-related myelodysplasia and acute myeloid leukemia after exposure to alkylating agents and significantly associated with deletion or loss of 5q, a complex karyotype, and a poor prognosis. *J. Clin. Oncol.* **19**, 1405–1413 (2001).
- Shih, A. H. et al. Mutational analysis of therapy-related myelodysplastic syndromes and acute myelogenous leukemia. *Haematologica* **98**, 908–912 (2013).
- Kayser, S. et al. The impact of therapy-related acute myeloid leukemia (AML) on outcome in 2853 adult patients with newly diagnosed AML. *Blood* **117**, 2137–2145 (2011).
- Link, D. C. et al. Identification of a novel TP53 cancer susceptibility mutation through whole-genome sequencing of a patient with therapy-related AML. *J. Am. Med. Assoc.* **305**, 1568–1576 (2011).
- The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
- Walter, M. J. et al. Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1090–1098 (2012).
- Walter, M. J. et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia* **27**, 1275–1282 (2013).
- Ding, L. et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
- Petitjean, A. et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum. Mutat.* **28**, 622–629 (2007).
- Welch, J. S. et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
- Catlin, S. N., Busque, L., Gale, R. E., Gutter, P. & Abkowitz, J. L. The replication rate of human hematopoietic stem cells *in vivo*. *Blood* **117**, 4460–4466 (2011).
- Carvajal, L. A. & Manfredi, J. J. Another fork in the road—life or death decisions by the tumour suppressor p53. *EMBO Rep.* **14**, 414–421 (2013).
- Chen, X., Ko, L. J., Jayaraman, L. & Prives, C. p53 levels, functional domains, and DNA damage determine the extent of the apoptotic response of tumor cells. *Genes Dev.* **10**, 2438–2451 (1996).
- Purvis, J. E. et al. p53 dynamics control cell fate. *Science* **336**, 1440–1444 (2012).
- Zhang, X. P., Liu, F., Cheng, Z. & Wang, W. Cell fate decision mediated by p53 pulses. *Proc. Natl Acad. Sci. USA* **106**, 12245–12250 (2009).
- Zhang, X. P., Liu, F. & Wang, W. Two-phase dynamics of p53 in the DNA damage response. *Proc. Natl Acad. Sci. USA* **108**, 8990–8995 (2011).
- Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253**, 49–53 (1991).
- Kandoth, C. et al. Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
- Bondar, T. & Medzhitov, R. p53-mediated hematopoietic stem and progenitor cell competition. *Cell Stem Cell* **6**, 309–322 (2010).
- Marusyk, A., Porter, C. C., Zaberezhnyy, V. & DeGregori, J. Irradiation selects for p53-deficient hematopoietic progenitors. *PLoS Biol.* **8**, e1000324 (2010).
- Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
- Birch, J. M. et al. Relative frequency and morphology of cancers in carriers of germline TP53 mutations. *Oncogene* **20**, 4621–4628 (2001).
- Casás-Selves, M. & Degregori, J. How cancer shapes evolution, and how evolution shapes cancer. *Evolution* **4**, 624–634 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Schmidt, B. McKethan and R. Miller for technical assistance, and K. Odell and J. Tucker-Davis for animal care. We thank P. Goodfellow and J. Ivanovich for providing peripheral blood leukocyte genomic DNA from cancer-free individuals. This work was supported by National Institutes of Health grants P01 CA101937 (D.C.L.) and U54 HG003079 (R.K.W.) and by a grant from the Leukemia & Lymphoma Society (D.C.L.).

Author Contributions T.N.W. and G.R. designed and performed the research, analysed the data, and wrote the manuscript. A.L.Y. and T.E.D. developed and optimized the amplicon-based random primer sequencing assay. C.A.M., D.S., J.H., R.S.F., L.D., E.R.M. and R.K.W. contributed to the generation and analysis of the whole-genome or targeted sequencing. W.T., T.L.L., S.H., J.M.K. and P.W. collected and processed clinical data and tissue samples. J.D.B. performed statistical analyses of the clinical data. J.S.W., J.F.D., M.J.W., T.A.G. and T.J.L. contributed to data analysis. D.C.L. supervised all of the research and edited the manuscript, which was approved by all co-authors.

Author Information Sequence information on the 22 t-AML whole-genome sequencing patients and one exome sequencing patient (UPN 967645) has been deposited in dbGaP under accession number phs000159. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.C.L. (dlink@dom.wustl.edu).

METHODS

Patient characteristics. For the whole-genome sequencing study, we intentionally selected the original 22 cases of t-AML to have minimal numbers of cytogenetic abnormalities. However, the additional 89 cases of t-AML/t-MDS were randomly selected from those samples with sufficient tumour and skin DNA. All patients were selected from a larger cohort of adult AML and MDS patients enrolled in a single institution tissue banking protocol that was approved by the Washington University Human Studies Committee (WU HSC#01-1014). Written informed consent for whole-genome sequencing was obtained from all study participants. Patients were treated in accordance with National Comprehensive Cancer Network (NCCN) guidelines (<http://www.nccn.org>) with an emphasis on enrolment in therapeutic clinical trials whenever possible. Clinical data for all patients, including the pre-existing condition requiring cytotoxic therapy, the cytotoxic therapy received before the t-AML/t-MDS diagnosis, cytogenetics, treatment approach and outcomes data, are presented in Extended Data Tables 1 and 2 and Supplementary Table 1. Peripheral blood leukocyte genomic DNA from cancer-free individuals (median age = 75.3 \pm 6.6 years) was obtained as part of a Washington University Institutional Review Board-approved protocol. All subjects had no previous history of invasive cancer or treatment with cytotoxic therapy, as determined by the medical history.

Whole-genome sequencing and variant detection. A previously described procedure²⁵ was followed for library construction and whole-genome sequencing. Briefly, Illumina DNA sequencing was used to generate sequence that covered the haploid reference at a depth between 30.51 and 72.60 (Supplementary Table 9). Sequence data were aligned to reference sequence build NCBI human build 36 using BWA v.0.5.5 (ref. 26) (params: -t 4) then merged and deduplicated using Picard v.1.29. We detected SNVs using the intersection of SAMtools v.r963 (ref. 27) (params: -A -B) and Somatic Sniper v.0.7.3 (ref. 28) (params: -q 1 -Q 15), and filtered to remove false positives (params: min-base-quality 15, min-mapping-quality 40, min-somatic-score 40). Indels were detected using GATK version 5336 (ref. 29) unioned with Pindel v.0.5 (ref. 30). Somatic copy number alterations were detected using copyCat v.1.5 (<http://github.com/chrisamiller/copycat>). We detected structural variants using Break-Dancer v.1.2 (ref. 31) and SquareDancer v.0.1 (<https://github.com/genome/genome/blob/master/lib/perl/Genome/Model/Tools/Sv/SquareDancer.pl>), followed by assembly with Tigr-SV (<https://github.com/genome/tigr-sv>). SciClone (in review; <http://github.com/genome/sciclon>) was used to infer the subclonal architecture of all whole-genome sequencing samples.

Validation and extension sequencing with variant detection. We used custom sequence capture arrays from Roche Nimblegen that targeted variants detected by whole-genome sequencing and extended this array to cover all coding exons from an additional 149 genes of interest (Supplementary Table 4). Libraries were prepared, sequence was generated, and somatic alterations identified as described for whole-genome sequencing with the addition of VarScan v.2.2.6 (ref. 32) (params: -min-var. -freq 0.08 -p-value 0.10 -somatic-p-value 0.01 -validation) as a variant caller for both SNVs and indels. On average, genes were covered with a depth of 58.3 (Supplementary Table 10). Biallelic *TP53* mutations in case 530447 were confirmed with PCR amplification of the genomic region containing both somatic mutations from the diagnostic t-AML sample. The resulting amplicons were cloned into the pCR-TOPO plasmid vector (Life Sciences) and sequenced using Sanger sequencing.

Statistical analyses. Fisher's exact tests were used to evaluate the association between pairs of dichotomous variables, with a significant right-sided *P* value indicating a positive relationship and a significant left-sided *P* value indicating a negative relationship. The relationship between overall survival and each discrete measure was tested with Kaplan–Meier survival analyses with separate analyses for the AML and MDS groups. Age at diagnosis was discretized into quartiles for each group. Multivariate proportional-hazards regression models were created separately for the AML and MDS groups. All variables with log-rank *P* values of 0.20 or less in the Kaplan–Meier analyses were included in the first step. In successive steps, the variable with the largest *P* value was removed and the model re-run until all remaining variables had *P* values of 0.05 or less. Two-way interactions among the remaining variables were examined. Variables removed in earlier steps were added back to the model one at a time to determine if they significantly improved the final model. The proportionality assumption was evaluated for each variable in the final models.

Rare variant detection using unique adaptor next-generation sequencing. Amplicons approximately 200 bp in length were prepared from patient genomic DNA samples using primers designed to amplify genomic regions harbouring known tumour-specific SNVs (Supplementary Table 11). These amplicons were prepared for next-generation sequencing (NGS) using the Illumina TruSeq DNA Sample Preparation Kit (Illumina Catalog #FC-121-2001) replacing the kit adaptors with adaptors containing a random nucleotide index sequence. Libraries were quantified using the Agilent qPCR NGS Library Quantification Kit, Illumina GA (Agilent Technologies Catalog #G4880A). Using this quantification, each library was diluted to ensure that each random index would be observed in multiple sequenced reads^{33,34}. Each diluted library was amplified and sequenced on the Illumina MiSeq platform.

Sequenced reads containing the same index sequence were grouped together creating 'read families' in a manner similar to established methods^{33,34}. Reads within a read family were aligned against each other to filter out stochastic sequencing errors generating an error-corrected read family consensus sequence. Each consensus sequence was locally aligned to UCSC hg19/GRCh37 using bowtie2 (ref. 35) with the default settings. The aligned read families were processed with Mpileup²⁷ using the parameters -BQ0 -d 1000000000000. Next, variants were called with VarScan³² using the parameters -min-coverage 10000 -min-reads2 10 -min-avg-qual 0 -min-var. -freq 0 -p-value 1. Variant allele frequencies for the expected mutations and the background error rate were visualized using IGV³⁶ and graphically represented using ggplot2 (ref. 37). Variant coordinates are displayed in hg18/GRCh36.

Detection of somatic *TP53* mutations in cancer-free subjects. Amplicons were prepared from healthy control genomic DNA samples using primers designed to amplify exons 4–8 of *TP53* (Supplementary Table 11). Patient-specific barcodes, 6 nucleotides in length, were appended to the 5' end of each primer to enable pooling of multiple samples for sequencing. Amplicons generated from each *TP53* exon/patient sample combination were generated as previously described and purified products were pooled in equimolar amounts. The pooled barcoded amplicons were prepared for error-corrected sequencing as previously described. Sequencing was completed on the Illumina Hi-Seq 2500 platform. Sequenced reads were demultiplexed based on the known patient-specific barcode sequences using a 2-nucleotide hamming distance. Demultiplexed sequence reads were organized into read families based on their random oligonucleotide index sequence and error-corrected as outlined previously. Read families composed of three reads or more were used for analysis. A binomial distribution of the substitution rate at each covered base in *TP53* was used to identify individuals with somatic *TP53* mutations. A variant was called if it met the following criteria: (1) the binomial *P* value was less than 10^{-6} ; (2) the VAF was greater than 1:10,000; (3) at least 10,000 unique read families were sequenced at the position of interest; (4) at least 10 read families called the variant; and (5) the VAF in the individual was greater than five times the mean VAF for all individuals with greater than 10,000 \times coverage at that specific nucleotide. Read families from one patient sample (barcode GTACGGC) were removed from analysis due to a high error rate. All somatic mutations were identified in this manner except for *TP53* Y220C, which received closer manual inspection due to the large number of these mutations observed in our t-AML cohort.

Extraction of genomic DNA from FFPE samples. Genomic DNA was extracted from FFPE samples with the QIAamp DNA FFPE Tissue Kit. Because of the effects of formalin fixation (cross-linking, DNA fragmentation, and so on), the amount of amplifiable DNA per sample was less than would be expected with Qubit fluorometric quantitation. As such, ddPCR was used to quantify the amount of amplifiable genomic DNA per sample such that the numbers of amplifiable domains tested were comparable between experimental and control samples.

ddPCR. All primers and probes for ddPCR were designed by Bio-Rad as per MIQE guidelines³⁸. In the case of *TP53* Y220C, the *TP53* region of interest in exon 6 was amplified with the following primers: 5'-TTTTCGACATAGTGTGGTG-3' and 5'-CTGACAACCAACCTTAAC-3'. The 5'-Hex/TGCCCTATGAGCCGCT/Iowa Black FQ-3' probe was used to detect the wild-type allele and the 5'-FAM/CCCTGTGAGCCGCTGA/Iowa Black FQ-3' probe was used to detect the mutant allele. All reagents were purchased from Bio-Rad. ddPCR was performed as previously described³⁹. Specifically, quantitative PCR was performed with 900–1,800 nM forward and reverse primers, 250 nM mutant and wild-type genomic probes, and 2–4 ng μL^{-1} genomic DNA. Quantitative PCR was performed with annealing/extension temperatures of 55.5–60 °C for 40 cycles. For droplet generation and analysis, we used the Bio-Rad QX100 and QX200 Droplet Digital PCR Systems.

Owing to the fact that DNA degradation with time (that is, guanosine oxidation, cytosine deamination) is known to interfere with rare allele detection³⁴, we only identified variant alleles present in droplets also lacking the reference allele. This greatly increased the specificity of our calls by removing droplets in which one of the two DNA strands may have been chemically altered. At low variant allele frequency, it was assumed that only a single variant allele was present in these 'mutant only' drops. Droplet allele distribution follows a Poisson distribution such that the number of droplets only containing a single allele (either variant or reference) can be determined from the percentage of empty droplets. Of note, droplets showing evidence of template independent amplification (that is, observed in 'no template controls') were counted as empty droplets. The VAF was determined from the fraction of the single allele droplets containing the variant allele. When appropriate, control samples were used to subtract potential background signal. VAFs calculated in this method were highly concordant with VAFs obtained through unique-adaptor NGS.

Generation and analysis of *TP53*^{+/-} bone marrow chimaeras. *TP53*^{+/-} and wild-type mice were inbred on a C57BL/6 strain. Bone marrow from *TP53*^{+/-} mice expressing *Ly5.2* was mixed at a 1:7 ratio with bone marrow from wild-type mice expressing *Ly5.1* and transplanted retro-orbitally into lethally irradiated *Ly5.1/5.2*

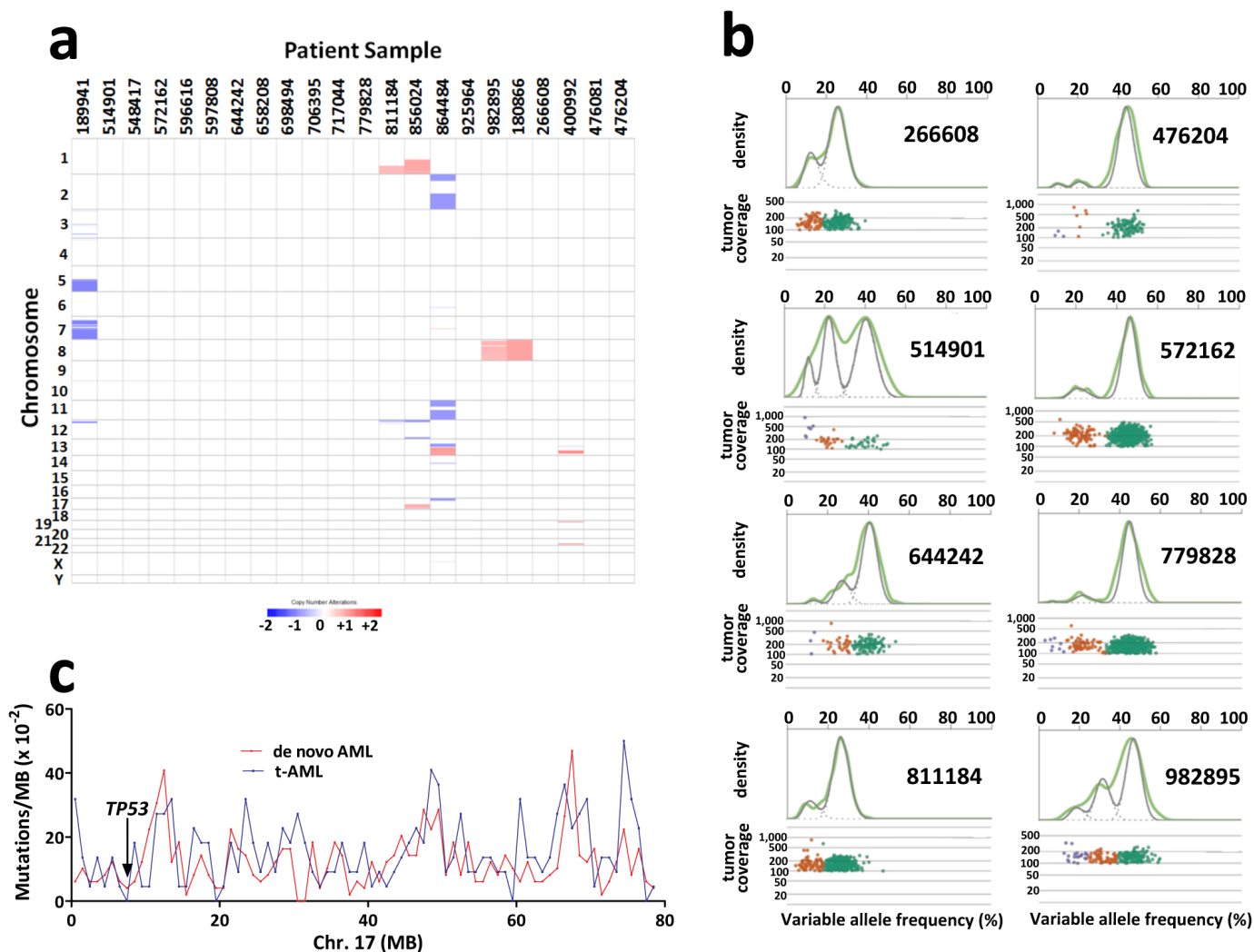
recipients. $Tp53^{+/-}$ and wild-type donors were both age (6–12 weeks) and sex matched (female). A total of 3×10^6 cells were injected per recipient mice. Recipient mice were conditioned with 1,000–1,100 cGy from a 137 caesium source at a rate of approximately 95 cGy min^{-1} before transplantation. Prophylactic antibiotics (trimethoprim-sulfamethoxazole; Alparma) were given during the initial 2 weeks after transplantation. Five weeks after transplantation, mice were given two doses of ENU (100 mg kg^{-1} ; Sigma-Aldrich) or vehicle alone intraperitoneally 9 days apart. Mice were stratified according to $Tp53^{+/-}$ chimaerism and then randomly distributed into the ENU and vehicle controls such that both cohorts had similar levels of $Tp53^{+/-}$ chimaerism at baseline. ENU and placebo were delivered in a final solution with 10% DMSO, 90 mM sodium citrate, and 180 mM sodium phosphate, pH 5.0. Peripheral blood chimaerism was measured before ENU administration and 4–12 weeks after ENU administration. The investigator was not blinded. Mice were euthanized and bone marrow chimaerism analysed 12 weeks after ENU administration. The desired cohort size was determined based on observations from previously reported experiments²⁰, and two independent experiments were performed. Mice were maintained under standard pathogen-free conditions according to methods approved by the Washington University animal studies committee.

Flow cytometry. Flow cytometry data were collected on a Gallios 10-colour, 3-laser flow cytometer (Beckman Coulter) and analysed with FlowJo software (Treestar). Cells were stained by standard protocols with the following antibodies (eBiosciences unless otherwise noted): Ly5.1 (A20, CD45.1), Ly5.2 (104, CD45.2), Ly6C/G (RB6-8C5, Gr-1), CD3e (145-2C11), CD45R (RA3-6B2, B220), CD11c (N418), TER-119, CD41 (MWReg30), CD117 (ACK2, c-Kit) and Ly-6A/E (D7, Sca).

Estimation of $TP53$ mutation frequency in ageing stem cells. The frequency and profile of somatic single-nucleotide mutations in the HSCs of normal individuals have been previously measured¹¹. The somatic mutational burden is ageing-related, and the estimated rate of mutagenesis obtained from this study was 3.2×10^{-9} mutations per nucleotide per year (95% confidence interval $2.4\text{--}4.0 \times 10^{-9}$) for the average nucleotide in the exome. Thus, we would predict an average 50 year old to have 1.6×10^{-7} mutations per position. These mutations would not be randomly distributed but biased (in particular towards C to T/G to A transitions). It has been previously proposed that an individual possesses approximately 10,000 distinct HSCs¹². We used a randomized Monte Carlo simulation to model the prevalence of somatic single-nucleotide mutations in healthy 50 year olds with 10,000 HSCs given a normal somatic mutational profile and mutation rate. Repeated simulation ($n = 100,000$) allowed us to predict the distribution of ageing-induced $TP53$

(NM_000546) somatic mutations. As expected, this simulation modelled a Poisson process. We classified $TP53$ mutations as likely to be functional if they fulfilled both of the following criteria. First, we analysed the mutations using the SIFT program (<http://sift.jcvi.org>) and required a SIFT score ≤ 0.05 . Second, we required that the somatic mutations be reported at least once if a nonsense mutation or at least twice if a missense mutation in the International Agency for Research on Cancer $TP53$ database (<http://p53.iarc.fr>). On the basis of this simulation, we predict that 44% of 50-year-old individuals harbour one or more HSCs with a functional $TP53$ mutation.

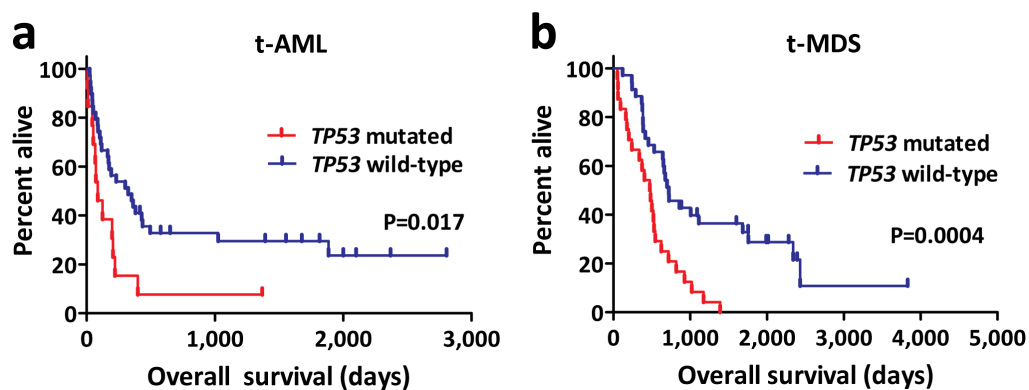
25. Mardis, E. R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361**, 1058–1066 (2009).
26. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
27. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
28. Larson, D. E. *et al.* SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311–317 (2012).
29. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
30. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).
31. Chen, K. *et al.* BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nature Methods* **6**, 677–681 (2009).
32. Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
33. Kinde, I., Wu, J., Papadopoulos, N., Kinzler, K. W. & Vogelstein, B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl Acad. Sci. USA* **108**, 9530–9535 (2011).
34. Schmitt, M. W. *et al.* Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl Acad. Sci. USA* **109**, 14508–14513 (2012).
35. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
36. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
37. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009).
38. Bustin, S. A. *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).
39. Hindson, B. J. *et al.* High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* **83**, 8604–8610 (2011).



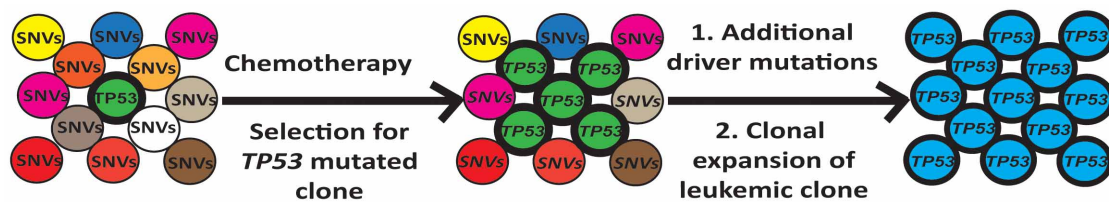
Extended Data Figure 1 | Whole-genome sequencing analysis of t-AML.

a, Somatic copy number alterations in the 22 cases of t-AML. Blue indicates copy number loss; red indicates copy number gain. **b**, Representative clonality plots for 8 cases of t-AML are shown. Scatter plots (bottom) show variant allele frequency and read depth in the tumour sample. Variant alleles in the founding clone are depicted in green, while variants in subclones are depicted in

orange or purple. Top, kernel density plots of the VAF data (green line) along with the posterior predictive densities (grey line) from the mathematical model used to segregate clusters. **c**, Frequency of tier 1 silent, tier 2, and tier 3 mutations in 1 Mb increments across chromosome 17 in *de novo* AML and t-AML. The *TP53* genomic locus is identified.

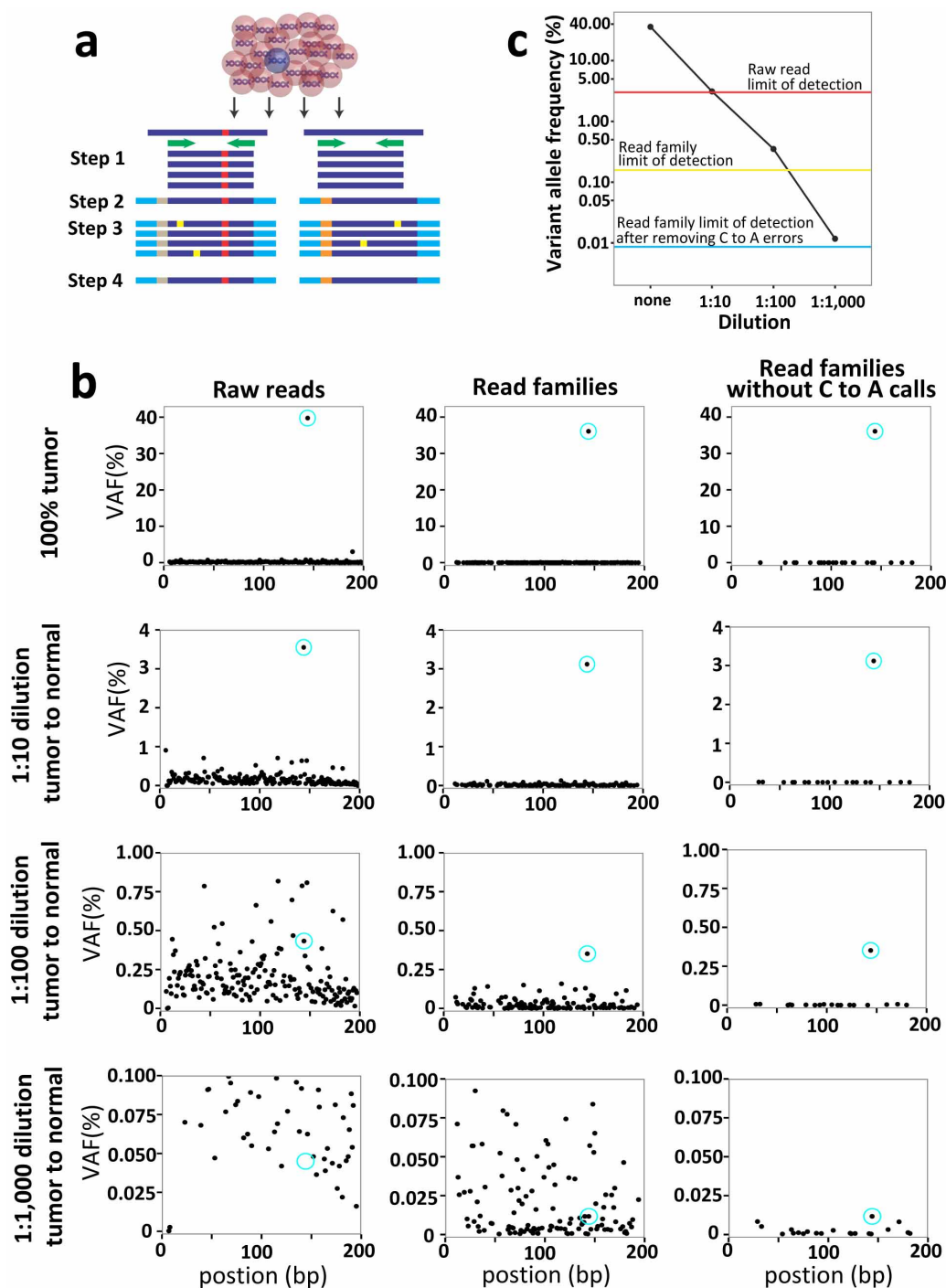


Extended Data Figure 2 | *TP53* mutations are associated with decreased overall survival in t-AML/t-MDS. **a**, Overall survival in *TP53* mutated ($n = 13$) and *TP53* wild-type ($n = 39$) t-AML patients. **b**, Overall survival in *TP53* mutated ($n = 24$) and *TP53* wild-type ($n = 35$) t-MDS patients.



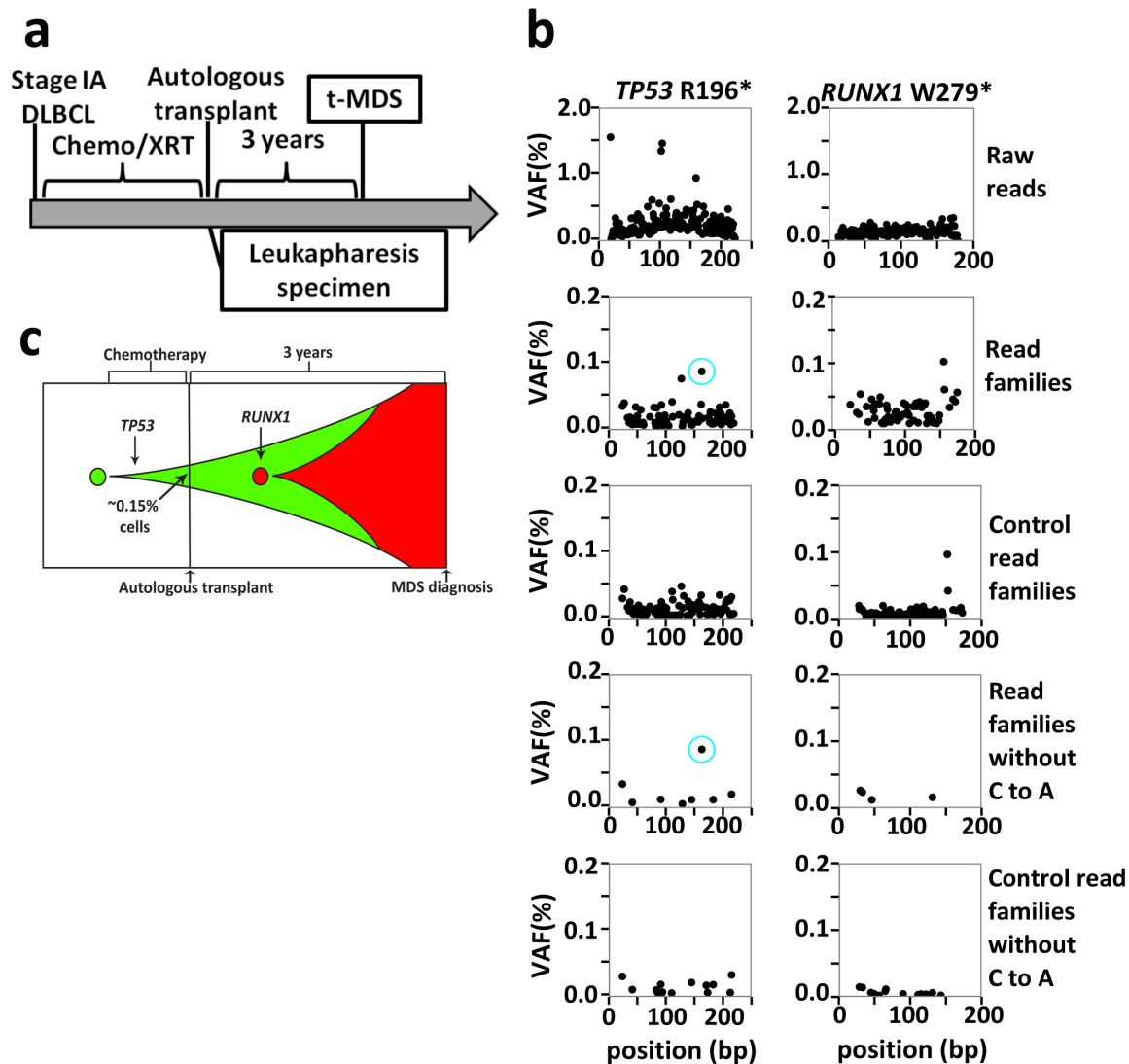
Extended Data Figure 3 | Model of how cytotoxic therapy shapes clonal evolution in t-AML/t-MDS. Age-related mutations in HSPCs result in the production of a genetically heterogeneous population of HSPCs, including rare HSPCs with heterozygous *TP53* mutations in some individuals. During chemotherapy and/or radiotherapy for the primary cancer, HSPC clones

harbouring a *TP53* mutation have a competitive advantage, resulting in expansion of that clone. Subsequent acquisition of additional driver mutations results in transformation to t-AML/t-MDS. Of note, the presence of *TP53* mutations probably accounts for the high incidence of cytogenetic abnormalities in t-AML/t-MDS and poor response to chemotherapy.



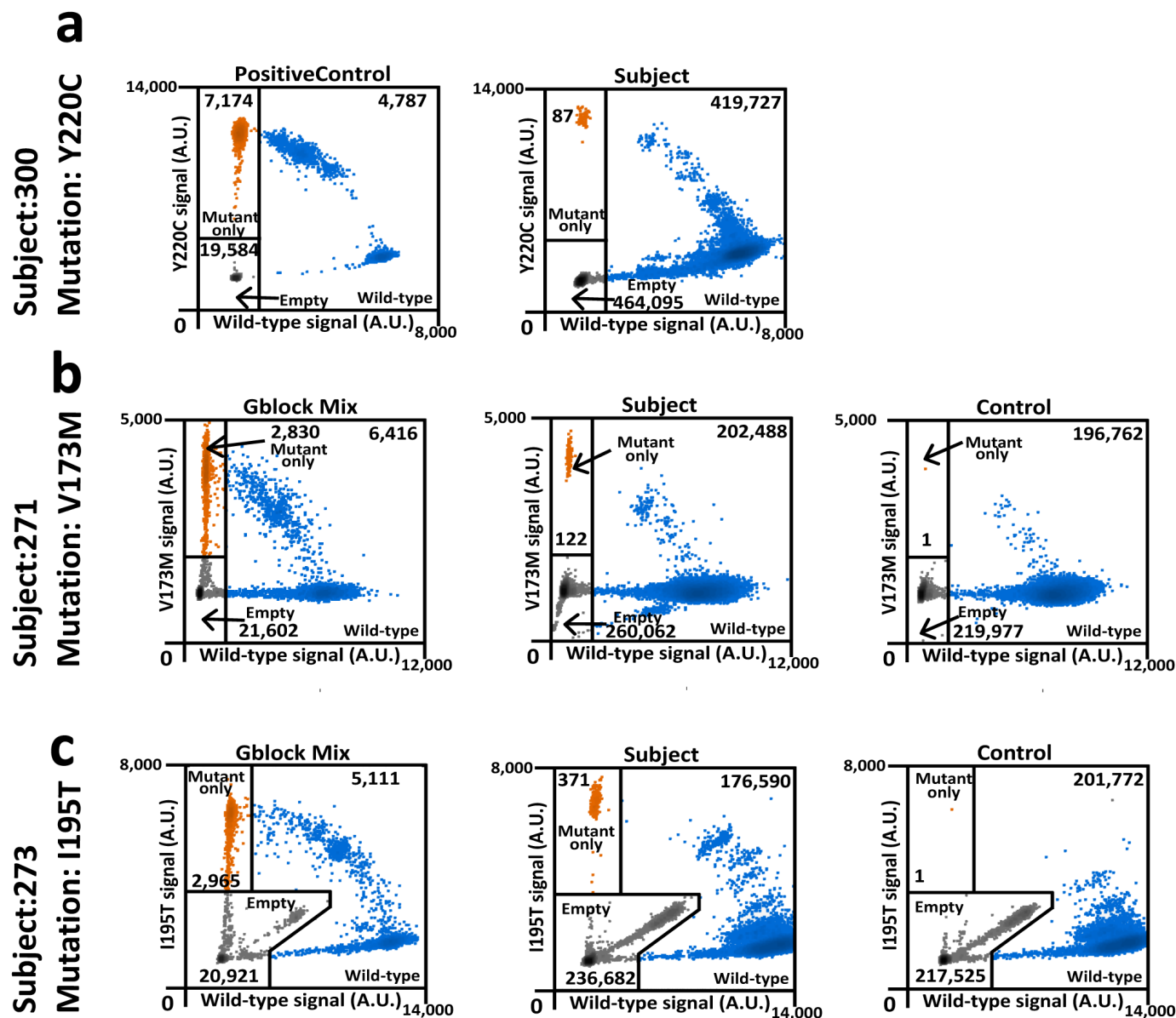
Extended Data Figure 4 | Validation of the unique adaptor sequencing method. **a**, Unique adaptor sequencing approach. Step 1: genomic DNA is amplified with *TP53*-specific primers (green) with subpopulation-specific variant alleles highlighted in red. Step 2: randomly indexed adapters (tan and grey) are ligated to each amplicon. Step 3: the indexed amplicons are amplified to generate multiple reads possessing the same barcode (that is, read families). Step 4: after sequencing, reads are aligned and grouped by read families to generate an error-corrected consensus sequence. Sequencing errors (yellow) are randomly distributed amongst read families, while true variant alleles (red) are present in all members of a given read family. **b**, A tumour sample (UPN 895681) with a known *TP53* somatic mutation (chromosome

17: 7519119 T to A) at a VAF of ~37% was mixed with normal genomic DNA sample at the indicated ratio, and conventional (left) or unique adaptor next-generation sequencing (middle and right) was performed, as described in Methods. DNA degradation with time may result in errors that are then amplified during PCR, providing a source of false-positive calls. This is particularly true for C to A transversions. Since none of the *TP53* mutations analysed in this study were C to A transversions, we also analysed the data after removing C to A calls (right). The *TP53* variant allele is circled in blue. **c**, The threshold of detection for the variant allele with each sequencing method is shown.



Extended Data Figure 5 | Clonal evolution in case 314666. **a**, Clinical course of case 314666. Chemo, chemotherapy; DLBCL, diffuse large B-cell lymphoma; XRT, radiotherapy. **b**, Unique adaptor sequencing was performed on genomic DNA derived from leukapheresis samples obtained 3 years before the

diagnosis of t-MDS for the two clonal mutations present in the diagnostic t-MDS sample. Genomic DNA from a patient lacking these variants was used as a control. The blue circle indicates the position of the variant SNV. **c**, Proposed model of clonal evolution to t-MDS in this case.



Extended Data Figure 6 | ddPCR verification of selected somatic *TP53* mutations identified in peripheral blood of cancer-free individuals.
a–c, ddPCR was performed on genomic DNA isolated from the peripheral blood of cancer-free individuals (middle) for whom unique-read adaptor sequencing suggested the presence of the indicated *TP53* mutation. Controls represent peripheral blood DNA from cancer-free elderly individuals with VAFs not above background levels for the mutation of interest (right); the negative control for *TP53* Y220C is shown in Fig. 3b. a, The diagnostic t-AML

sample from patient 967645 was used as a positive control for *TP53* Y220C. b, c, For *TP53* V173M (b) and *TP53* I195T (c) double-stranded genomic blocks (gBlocks) were synthesized containing the mutation of interest and mixed with gBlocks of wild-type sequence. Droplets containing only the variant *TP53* allele are highlighted in orange, droplets containing the wild-type *TP53* allele (with or without the variant *TP53* allele) are highlighted in blue; empty droplets are grey. The number of droplets in each gate is indicated.

Extended Data Table 1 | Clinical summary of the 22 t-AML whole-genome sequencing cases

Age	Median	56.5 years (26-80)
Gender	Male	36.4%
	Female	63.6%
Prior Disease	Breast	45.5%
	Non-Hodgkin's Lymphoma	13.6%
	Multiple Sclerosis	9.1%
Known Previous Treatment	Other	31.8%
	Alkylator	45.5%
	Topoisomerase inhibitor	63.6%
	Radiation	68.2%
	Autologous Transplant	9.1%
Latency	Median	3.2 years (0.9-13.3)
Cytogenetics	complex	22.7%
	MLL rearrangement	22.7%
	non-complex non-MLL	63.6%
% Blasts in the bone marrow	Median	79% (19-95%)
Most intensive t-AML/t-MDS treatment regimen	Allogeneic transplant	27.3%
	Myeloablative	40.9%
	Non-myeloablative	13.6%
	Other/unknown	18.2%
Remission	Yes	50%
	No	40.9%
Overall Survival	Median	140.5 days (8-2000)

Latency is defined as the time from the original cancer diagnosis to the development of t-AML/t-MDS.

Extended Data Table 2 | Clinical summary of the combined 111 t-AML/t-MDS cases

Age	Median	61 years (18-82)
Gender	Male	53.2%
	Female	46.8%
Prior Disease	Breast	29.7%
	Non-Hodgkin's Lymphoma	24.7%
	Hodgkin's Disease	22.5%
	Other	37.8%
Known Previous Treatment	Alkylator	55.9%
	Topoisomerase inhibitor	50.5%
	Radiation	63.1%
	Autologous Transplant	21.6%
Latency	Median	6.25 years (0.4-40.7)
Diagnosis	Leukemia	46.8%
	MDS	53.2%
Cytogenetics	deletion 5	26.1%
	deletion 7	28.8%
	complex	45.0%
	MLL rearrangement	5.4%
	other/unknown	41.4%
% Blasts in the bone marrow	Median	13% (0-95%)
Most intensive AML treatment regimen	Allogeneic transplant	41.4%
	Myeloablative	17.1%
	Non-myeloablative	24.3%
	Other/unknown	17.1%
Remission	Yes	49.5%
	No	43.2%
Overall Survival	Median	414 days (8-3831)

Latency is defined as the time from the original cancer diagnosis to the development of t-AML/t-MDS.

Extended Data Table 3 | Previously banked tissue samples in patients with t-AML/t-MDS with clonal *TP53* mutations

Patient	<i>TP53</i> mutation in t-AML/t-MDS			Prior banked tissue sample		
	Position (Chr 17)	Mutation	Coding change	Year of Banking	Year of Diagnosis	Prior Banked Tissue
236041	7,518,261	T to A	R249W	2007	2011	BM FFPE
341666	7,518,988	G to A	R196*	2002	2005	Pheresis
530447	7,519,238	C to G	K139N	2001	2007	Pheresis
530447	7,518,263	C to T	R248Q	2001	2007	Pheresis
648904	7,514,759	C to T	Exon 9 splice site	2001	2004	Pheresis
756582	7,519,015	C to T	Exon 6 splice site	1999	2007	Pheresis
895681	7,519,119	T to A	H179L	2000	2006	BM FFPE
967645	7,518,915	T to C	Y220C	2005	2010	BM Flow

All patients had one or more clonal *TP53* mutations in their diagnostic t-AML/t-MDS samples (530447 had biallelic mutations). Cases in which the previously banked sample had detectable *TP53* mutated cells are highlighted in red. See Supplementary Table 1 for the clinical and molecular features of these cases. BM FFPE, formalin-fixed paraffin-embedded sample; BM flow, snap-frozen bone marrow leukocyte pellet.

Extended Data Table 4 | Somatic *TP53* mutations in 19 cancer-free individuals

Sample	Chr	Exon	Start	Stop	Ref	Var	Amino acid	COSMIC ID	Var count	Total read family count	VAF (read-family)	VAF (ddPCR)
34	17	7	7518230	7518230	T	G	D259A	none	13	33085	0.039%	N.D.
99	17	7	7518273	7518273	C	T	G245S	COSM6932	18	41836	0.043%	N.D.
99	17	8	7517849	7517849	C	T	V272M	COSM10891	26	81015	0.032%	N.D.
269	17	8	7517845	7517845	C	T	R273H	COSM10660	489	420026	0.12%	N.D.
271	17	5	7519138	7519138	C	T	V173M	COSM11084	177	182809	0.097%	0.081%
271	17	5	7519174	7519174	C	T	A161T	COSM10739	25	164591	0.015%	N.D.
271	17	NA	7520035	7520035	A	T	SPLICING	COSM1522474	23	165672	0.014%	N.D.
271	17	NA	7517934	7517934	C	T	INTRONIC	none	36	333996	0.011%	N.D.
273	17	6	7518990	7518990	A	G	I195T	COSM11089	57	15540	0.37%	0.28%
300	17	6	7518915	7518915	T	C	Y220C	COSM10758	91	316765	0.029%	0.029%
324	17	8	7517819	7517819	G	A	R282W	COSM10704	51	86090	0.059%	N.D.
335	17	7	7518264	7518264	G	C	R248G	COSM11564	245	218077	0.11%	N.D.
338	17	7	7518264	7518264	G	A	R248W	COSM10656	188	51001	0.37%	N.D.

Coverage statistics are as follows. In the amplicon targeting exon 4, 17/19 subjects had >10,000 coverage in 100% of the amplicon. In the amplicon targeting exon 5, 17/19 subjects had >10,000 coverage in 100% of the amplicon. In the amplicon targeting exon 6, 5/19 subjects had >10,000 coverage in 100% of the amplicon, and 11/19 subjects had >10,000 coverage in at least 75% of the amplicon. In the amplicon targeting exon 7, 17/19 subjects had >10,000 coverage in 100% of the amplicon. In the amplicon targeting exon 8, 18/19 subjects had >10,000 coverage in 100% of the amplicon. See Supplementary Table 11 for the primers used to make the amplicons from genomic DNA. N.D., not determined.

Enhancer–core–promoter specificity separates developmental and housekeeping gene regulation

Muhammad A. Zabidi^{1*}, Cosmas D. Arnold^{1*}, Katharina Schernhuber¹, Michaela Pagani¹, Martina Rath¹, Olga Frank¹ & Alexander Stark¹

Gene transcription in animals involves the assembly of RNA polymerase II at core promoters and its cell-type-specific activation by enhancers that can be located more distally¹. However, how ubiquitous expression of housekeeping genes is achieved has been less clear. In particular, it is unknown whether ubiquitously active enhancers exist and how developmental and housekeeping gene regulation is separated. An attractive hypothesis is that different core promoters might exhibit an intrinsic specificity to certain enhancers^{2–6}. This is conceivable, as various core promoter sequence elements are differentially distributed between genes of different functions⁷, including elements that are predominantly found at either developmentally regulated or at housekeeping genes^{8–10}. Here we show that thousands of enhancers in *Drosophila melanogaster* S2 and ovarian somatic cells (OSCs) exhibit a marked specificity to one of two core promoters—one derived from a ubiquitously expressed ribosomal protein gene and another from a developmentally regulated transcription factor—and confirm the existence of these two classes for five additional core promoters from genes with diverse functions. Housekeeping enhancers are active across the two cell types, while developmental enhancers exhibit strong cell-type specificity. Both enhancer classes differ in their genomic distribution, the functions of neighbouring genes, and the core promoter elements of these neighbouring genes. In addition, we identify two transcription factors—Dref and Trl—that bind and activate housekeeping versus developmental enhancers, respectively. Our results provide evidence for a sequence-encoded enhancer–core-promoter specificity that separates developmental and housekeeping gene regulatory programs for thousands of enhancers and their target genes across the entire genome.

We chose the core promoter of *Ribosomal protein gene 12* (*RpS12*) and a synthetic core promoter derived from the *even skipped* transcription factor¹¹ as representative ‘housekeeping’ and ‘developmental’ core promoters, respectively (hereafter termed hkCP and dCP; Fig. 1a and Extended Data Figs 1, 2) and tested the ability of all candidate enhancers genome wide to activate transcription from these core promoters using self-transcribing active regulatory region sequencing (STARR-seq)¹² in *D. melanogaster* S2 cells. This set-up allows the testing of all candidates in a defined sequence environment, which differs only in the core promoter sequences but is otherwise constant^{12,13}.

Two hkCP STARR-seq replicates were highly similar (genome-wide Pearson correlation coefficient (PCC) 0.98; Extended Data Fig. 1c) and yielded 5,956 enhancers, compared with 5,408 enhancers obtained when we reanalysed dCP STARR-seq data¹² (Supplementary Table 1). Interestingly, the hkCP and dCP enhancers were largely non-overlapping (Fig. 1b, c) and the genome-wide enhancer activity profiles differed (PCC 0.38), as did the individual enhancer strengths: of the 11,364 enhancers, 8,144 (72%) activated one core promoter at least twofold more strongly than the other, a difference rarely seen in the replicate experiments for each of the core promoters (Fig. 1d). Indeed, 21 out of 24 hkCP-specific enhancers activated luciferase expression (>1.5-fold and *t*-test $P < 0.05$) from the hkCP versus 1 out of 24 from the dCP (Fig. 1e and Extended Data Fig. 3). Consistently, 10 out of 12 dCP-specific enhancers were positive with the dCP but only 2 out of 12 with the hkCP, a highly significant difference ($P = 5.1 \times 10^{-6}$, Fischer’s exact test) that confirms the enhancer–core-promoter specificity observed for thousands of enhancers across the entire genome.

Enhancers that were specific to either the hkCP or the dCP showed markedly different genomic distributions (Fig. 2a and Extended Data

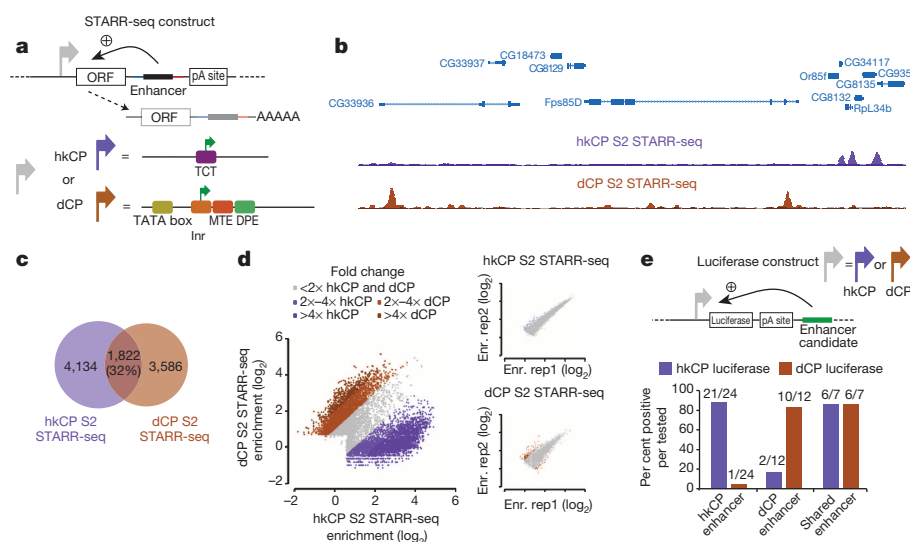


Figure 1 | Distinct sets of enhancers activate transcription from the hkCP and dCP in S2 cells. **a**, STARR-seq set-up using the hkCP housekeeping (*RpS12*; purple) and dCP developmental core promoters (*Drosophila* synthetic core promoter (DSCP)¹¹; brown). **b**, Genome browser screenshot depicting STARR-seq tracks for both core promoters. **c**, Overlap of hkCP and dCP enhancers. **d**, hkCP versus dCP STARR-seq enrichments at enhancers (insets show enrichment for replicates (Enr. rep1 versus 2 for hkCP and dCP; dCP data reanalysed from ref. 12)). **e**, hkCP, dCP or shared enhancers that activate luciferase (>1.5-fold and $P < 0.05$ (one-sided *t*-test); $n = 3$; Extended Data Figs 3 and 5) from hkCP (purple) or dCP (brown; numbers show positive/tested).

¹Research Institute of Molecular Pathology IMP, Vienna Biocenter VBC, Dr Bohr-Gasse 7, 1030 Vienna, Austria.

*These authors contributed equally to this work.

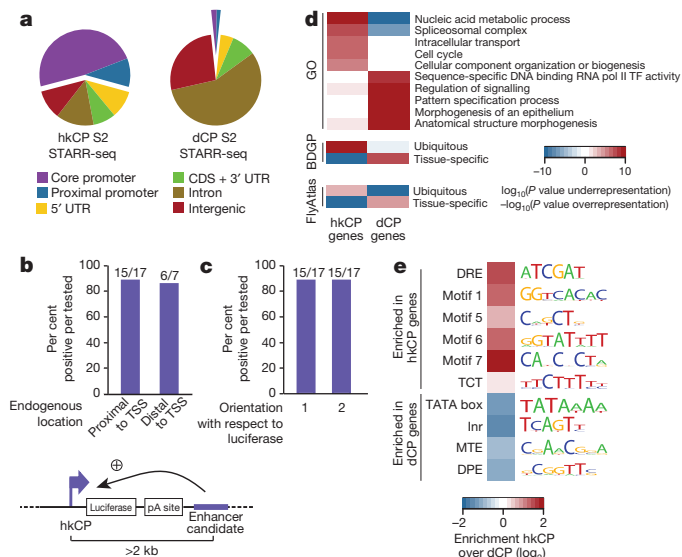


Figure 2 | hkCP and dCP enhancers differ in genomic distribution and flanking genes. **a**, Genomic distribution of hkCP and dCP enhancers. CDS, coding sequence; UTR, untranslated region. **b**, **c**, hkCP enhancers function distally in luciferase assays independent of their genomic positions (**b**) and orientation towards the luciferase TSS (**c**; orientation 1 from **b**; Extended Data Figs 3 and 5). **d**, **e**, GO (5 of the top 100 terms shown per column; Supplementary Table 11) and gene expression (terms curated from the Berkeley *Drosophila* Genome Project (BDGP) and FlyAtlas analyses (**d**) and enrichment of core promoter elements at TSSs (**e**) for genes next to hkCP and dCP enhancers. TF, transcription factor.

Fig. 4): whereas the majority (58.4%) of hkCP-specific enhancers overlapped with a transcription start site (TSS) or were proximal to a TSS (≤ 200 bp upstream; Fig. 2a), dCP-specific enhancers located predominantly to introns (56.5%) and intergenic regions (26.9%; Fig. 2a)¹². Importantly, despite the TSS-proximal location of most hkCP-specific enhancers, they activated transcription from a distal core promoter in STARR-seq (Fig. 1a and Extended Data Figs 1a, 2). Luciferase assays confirmed that they function from a distal position (>2 kb from the TSS) downstream of the luciferase gene and independently of their orientation towards the luciferase TSS (Fig. 2b, c and Extended Data Figs 3, 5). These results show that TSS-proximal sequences can act as bona fide enhancers¹⁴ and that developmental and housekeeping genes are both regulated through core promoters and enhancers, yet with a substantially different fraction of TSS-proximal enhancers (3.4% versus 58.4%).

hkCP and dCP enhancers were also located next to functionally distinct classes of genes according to gene ontology (GO) analyses: genes next to hkCP enhancers were enriched in diverse housekeeping functions

including metabolism, RNA processing and the cell cycle, whereas genes next to dCP enhancers were enriched for terms associated with developmental regulation and cell-type-specific functions (Fig. 2d, Extended Data Fig. 6a and Supplementary Tables 2–4). Consistently, hkCP enhancers were preferentially near ubiquitously expressed genes and dCP enhancers were near genes with tissue-specific expression (Fig. 2d and Supplementary Table 5).

The core promoters of the putative endogenous target genes of hkCP and dCP enhancers were also differentially enriched in known core promoter elements¹⁵ (Fig. 2e and Extended Data Fig. 6b): TSSs next to hkCP enhancers were enriched in Ohler motifs¹⁶ 1, 5, 6 and 7, consistent with the ubiquitous expression and housekeeping functions of these genes. In contrast, TSSs next to dCP enhancers were enriched in TATA box, initiator (Inr), motif ten element (MTE) and downstream promoter element (DPE) motifs, which are associated with cell-type-specific gene expression^{7,15}.

We next investigated whether the specificity that hkCP and dCP show to the two enhancer classes applies more generally. We selected three additional core promoters from housekeeping genes with different functions: from the *eukaryotic translation elongation factor 1 δ* (*eEF1 δ*), the putative splicing factor *x16*, and the cohesin loader *Nipped-B* (*NipB*). Importantly, all three contained combinations of core promoter elements that differed from that of hkCP, namely TCT⁸ and DNA-replication-related element (DRE) motifs (*eEF1 δ*), and Ohler motifs 1 and 6 (*x16* and *NipB*; Fig. 3a). In addition, we selected a DPE-containing core promoter of the transcription factor *pannier* (*pnr*) and the TATA-box core promoter of *Heat shock protein 70* (*Hsp70*), which can be activated by tissue-specific enhancers (for example, see ref. 17), thus covering the two most prominent core promoter types of regulated genes^{9,16,18}.

We performed STARR-seq for the five additional core promoters and grouped the genome-wide enhancer activity profiles of all seven core promoters by hierarchical clustering. This revealed two distinct clusters corresponding to the four housekeeping and the three developmental core promoters, respectively (Fig. 3b, Extended Data Fig. 7 and Supplementary Tables 6, 7), and the core promoters of both clusters indeed responded markedly differentially to individual genomic enhancers (Fig. 3c).

These results obtained for core promoters with diverse motif content and from genes with various functions suggest that the distinct enhancer preferences observed between hkCP and dCP apply more generally and that two broad classes of housekeeping and developmental (or regulated) core promoters exist. Differences within each class might correspond to differences in relative enhancer preferences of the core promoters^{2–6}, while similarities between both classes could reflect enhancers that are shared (Fig. 1c–e) or core promoters that can be activated to different extents by enhancers from both classes (for example,

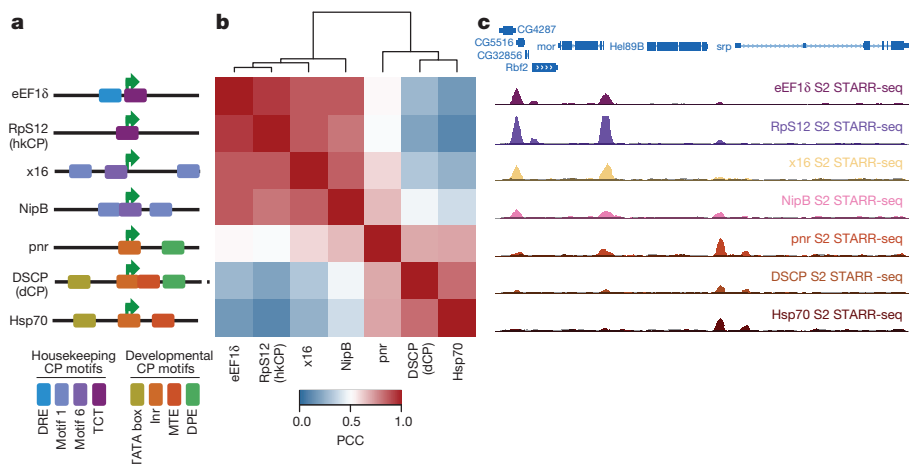


Figure 3 | Housekeeping and developmental core promoters differ characteristically in their enhancer preferences. **a**, Different housekeeping (top 4) and developmental-like (bottom 3) core promoters and their motif content (schematic). **b**, Bi-clustered heat map depicting pairwise similarities of STARR-seq signals (PCCs at peak summits). PCCs and dendrogram (top) show the separation between housekeeping and regulated core promoters. **c**, Genome browser screenshot depicting STARR-seq tracks for all seven core promoters.

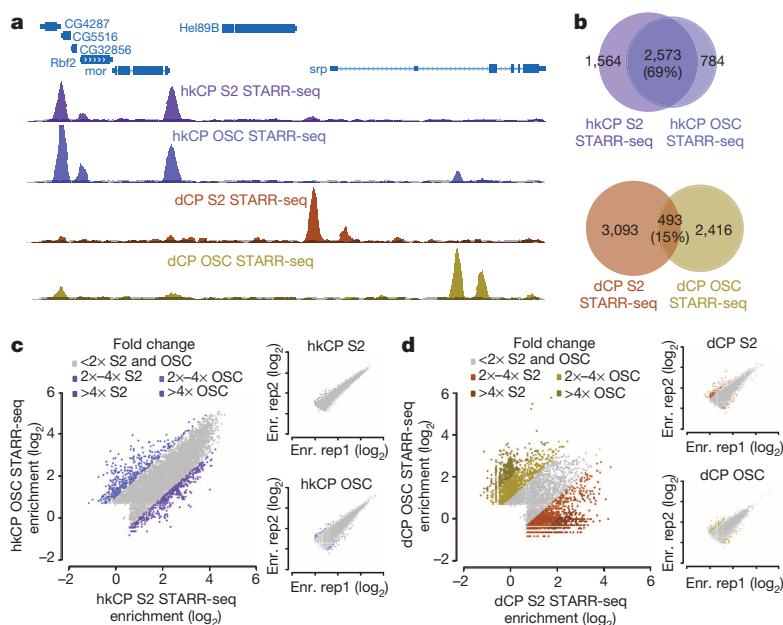


Figure 4 | hkCP enhancers are shared across cell types. **a**, Genome browser screenshot showing tracks for hkCP (top) and dCP STARR-seq (bottom) in S2 cells and OSCs. **b**, Overlap of hkCP (top) and dCP (bottom) enhancers between S2 cells and OSCs. **c**, **d**, hkCP (**c**) and dCP (**d**) STARR-seq enrichments in S2 cells versus OSCs at hkCP- or dCP-specific enhancers (insets show enrichments for replicates (Enr. rep) 1 versus 2; dCP data reanalysed from ref. 12).

NipB; Fig. 3b, c). The latter might be important if broadly expressed housekeeping genes need to be further activated in specific tissues.

To test whether hkCP enhancers function in different cell types, we performed STARR-seq using hkCP in OSCs, which differ strongly from S2 cells in gene expression and dCP enhancer activities¹². Two hkCP STARR-seq replicates in OSCs were highly similar (PCC 0.97) and yielded 6,217 enhancers (Supplementary Table 1), compared with 5,774 enhancers obtained for dCP data from OSCs¹². The OSC data confirmed the differences between hkCP and dCP enhancers observed in S2 cells (Extended Data Figs 8, 9 and Supplementary Tables 8–10). Strikingly,

the hkCP-specific enhancers in OSCs and S2 cells (3,357 and 4,137, respectively) were almost indistinguishable, whereas dCP-specific enhancers (2,909 in OSCs and 3,586 in S2 cells) differed strongly between the two cell types¹² and from the hkCP enhancers (Fig. 4a). The observation that hkCP enhancers showed similar activities in both cell types while dCP enhancers were cell-type specific was true genome wide when comparing genomic locations (69% versus 15% overlap) or enhancer strengths as measured by STARR-seq (PCC at peak summits 0.83 versus 0.05; Fig. 4b–d and Extended Data Fig. 9c). Together, these results show that hkCP enhancers are shared between two different cell types,

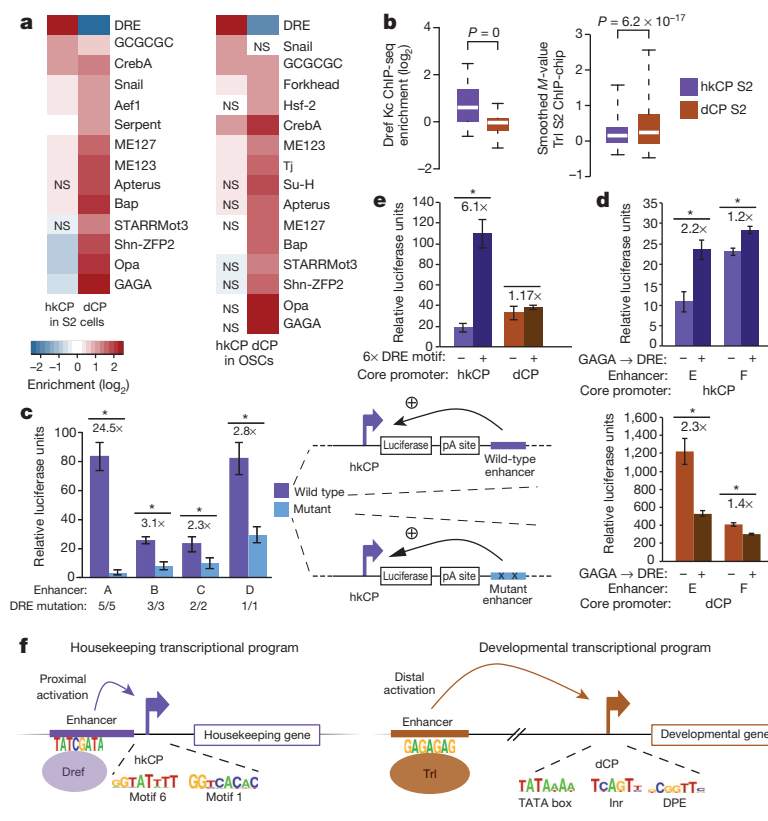


Figure 5 | hkCP and dCP enhancers depend on Dref and Trl, respectively. **a**, **b**, Motif enrichment (**a**) and ChIP signals for Dref and Trl (**b**) in hkCP and dCP enhancers. False discovery rate (FDR)-corrected hypergeometric $P > 0.01$; boxes: median and interquartile range; whiskers: 5th and 95th percentiles; two-sided Wilcoxon-rank-sum P values. NS, not significant. **c**, Luciferase assays for four wild-type and DRE-motif-mutant hkCP enhancers (numbers show mutated motifs). Error bars show standard deviation (s.d.) ($n = 3$, biological replicates). $*P < 0.005$ (one-sided t -test). **d**, Luciferase assays for two dCP enhancers (–) and their GAGA → DRE-mutant variants (+) with hkCP (top) and dCP (bottom; details as in c). **e**, Luciferase assays for an array of DRE motifs with hkCP and dCP (details as in c). **f**, Model: housekeeping genes contain Ohler motifs 1, 5, 6, 7 and/or the TCT motif and are activated by TSS-proximal hkCP enhancers via Dref. Regulated genes contain TATA box, Inr, MTE and/or DPE and are activated by distal dCP enhancers via Trl.

whereas dCP enhancers are cell-type specific¹², presumably representing ubiquitous housekeeping versus developmental and cell-type-specific gene expression programs.

To assess whether the marked core promoter specificities of the hkCP and dCP enhancers are encoded in their sequences, we analysed the *cis*-regulatory motif content of both classes of enhancers¹⁹. This revealed a strong enrichment of the DRE motif in hkCP enhancers (Fig. 5a and Supplementary Tables 11, 12), whereas dCP enhancers were strongly enriched in the GAGA motif of *Trithorax-like* (*Trl*) and other motifs previously described to be important for dCP enhancers²⁰. Published genome-wide chromatin immunoprecipitation (ChIP) data^{21,22} confirmed that DRE-binding factor (Dref) bound significantly more strongly to hkCP enhancers than to dCP enhancers (Wilcoxon $P = 0$; Fig. 5b), while the opposite was true for *Trl* (Wilcoxon $P = 6.2 \times 10^{-17}$). Considering only distal enhancers (>500 bp from the closest TSS) yielded the same results (Extended Data Fig. 10a, b and Supplementary Tables 13, 14), suggesting that the differential occupancy is a property of both classes of enhancers rather than a consequence of the different extents to which they overlap with TSSs. Disrupting the DRE motifs in four different hkCP enhancers substantially reduced the activities of the enhancers as measured by luciferase assays in S2 cells (between 2.3- and 24.5-fold reduction; Fig. 5c), while dCP enhancers depend on GAGA motifs²⁰. Adding DRE motifs to 11 different dCP enhancers significantly increased luciferase expression from the hkCP for 9 of them (82%; Extended Data Fig. 10c), and changing the GAGA motifs of two dCP enhancers to DRE motifs significantly increased the activities of both enhancers towards the hkCP but decreased their activities towards the dCP (Fig. 5d). Furthermore, an array of six DRE motifs was sufficient to activate luciferase expression from the hkCP but not the dCP (Fig. 5e). Together, these results show that hkCP and dCP enhancers depend on DRE and GAGA motifs, respectively, and demonstrate that DRE motifs are required and sufficient for hkCP enhancer function.

Our results show that developmental and housekeeping gene regulation is separated genome wide by sequence-encoded specificities of thousands of enhancers to one of two types of core promoter, supporting the longstanding 'enhancer-core-promoter specificity' hypothesis^{2-6,23}. Our findings indicate that these specificities are probably mediated by defined biochemical compatibilities²⁴ between different *trans*-acting factors such as Dref versus *Trl* (at enhancers) and the different paralogues that exist for several components of the general transcription apparatus (at core promoters), presumably including the TATA-box-binding protein-related factor 2 (*Trf2*) at housekeeping core promoters^{25,26}. As such paralogues can have tissue-specific expression and stage-specific or promoter-selective functions^{27,28} (reviewed in refs 29, 30), sequence-encoded enhancer-core-promoter specificities could be used more widely to define and separate different transcriptional programs (Fig. 5f).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 May; accepted 20 October 2014.

Published online 15 December 2014; corrected online 25 February 2015 (see full-text HTML version for details).

- Levine, M., Cattoglio, C. & Tjian, R. Looping back to leap forward: transcription enters a new era. *Cell* **157**, 13–25 (2014).
- Li, X. & Noll, M. Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the *Drosophila* embryo. *EMBO J.* **13**, 400–406 (1994).
- Ohtsuki, S., Levine, M. & Cai, H. N. Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev.* **12**, 547–556 (1998).
- Sharpe, J., Nonchev, S., Gould, A., Whiting, J. & Krumlauf, R. Selectivity, sharing and competitive interactions in the regulation of *Hoxb* genes. *EMBO J.* **17**, 1788–1798 (1998).
- Merli, C., Bergstrom, D. E., Cygan, J. A. & Blackman, R. K. Promoter specificity mediates the independent regulation of neighboring genes. *Genes Dev.* **10**, 1260–1270 (1996).
- Butler, J. E. & Kadonaga, J. T. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev.* **15**, 2515–2519 (2001).

- Kadonaga, J. T. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 40–51 (2012).
- Parry, T. J. et al. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev.* **24**, 2013–2018 (2010).
- Engström, P. G., Ho Sui, S. J., Drivenes, O., Becker, T. S. & Lenhard, B. Genomic regulatory blocks underlie extensive microsynteny conservation in insects. *Genome Res.* **17**, 1898–1908 (2007).
- FitzGerald, P. C., Sturgill, D., Shyakhnenko, A., Oliver, B. & Vinson, C. Comparative genomics of *Drosophila* and human core promoters. *Genome Biol.* **7**, R53 (2006).
- Pfeiffer, B. D. et al. Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl Acad. Sci. USA* **105**, 9715–9720 (2008).
- Arnold, C. D. et al. Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074–1077 (2013).
- Shlyueva, D. et al. Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin. *Mol. Cell* **54**, 180–192 (2014).
- Banerji, J., Rusconi, S. & Schaffner, W. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
- Lenhard, B., Sandelin, A. & Carninci, P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nature Rev. Genet.* **13**, 233–245 (2012).
- Ohler, U., Liao, G.-C., Niemann, H. & Rubin, G. M. Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol.* **3**, research0087.1–0087.12 (2002).
- Smith, D., Wohlgemuth, J., Calvi, B. R., Franklin, I. & Gelbart, W. M. *hobo* enhancer trapping mutagenesis in *Drosophila* reveals an insertion specificity different from *P* elements. *Genetics* **135**, 1063–1076 (1993).
- Kutach, A. K. & Kadonaga, J. T. The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Mol. Cell Biol.* **20**, 4754–4764 (2000).
- Yáñez-Cuna, J. O., Dinh, H. Q., Kvon, E. Z., Shlyueva, D. & Stark, A. Uncovering *cis*-regulatory sequence requirements for context-specific transcription factor binding. *Genome Res.* **22**, 2018–2030 (2012).
- Yáñez-Cuna, J. O. et al. Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features. *Genome Res.* **24**, 1147–1156 (2014).
- Gurudatta, B. V., Yang, J., Van Bortle, K., Donlin-Asp, P. G. & Corces, V. G. Dynamic changes in the genomic localization of DNA replication-related element binding factor during the cell cycle. *Cell Cycle* **12**, 1605–1615 (2013).
- modENCODE Consortium Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **330**, 1787–1797 (2010).
- Ohler, U. & Wassarman, D. A. Promoting developmental transcription. *Development* **137**, 15–26 (2010).
- van Arensbergen, J., van Steensel, B. & Bussemaker, H. J. In search of the determinants of enhancer-promoter interaction specificity. *Trends Cell Biol.* <http://dx.doi.org/10.1016/j.tcb.2014.07.004> (2014).
- Wang, Y.-L. et al. TRF2, but not TBP, mediates the transcription of ribosomal protein genes. *Genes Dev.* **28**, 1550–1555 (2014).
- Isogai, Y., Koles, S., Prestel, M., Hochheimer, A. & Tjian, R. Transcription of histone gene cluster by differential core-promoter factors. *Genes Dev.* **21**, 2936–2949 (2007).
- Hochheimer, A., Zhou, S., Zheng, S., Holmes, M. C. & Tjian, R. TRF2 associates with DREF and directs promoter-selective gene expression in *Drosophila*. *Nature* **420**, 439–445 (2002).
- Deato, M. D. E. & Tjian, R. Switching of the core transcription machinery during myogenesis. *Genes Dev.* **21**, 2137–2149 (2007).
- D'Alessio, J. A., Wright, K. J. & Tjian, R. Shifting players and paradigms in cell-specific transcription. *Mol. Cell* **36**, 924–931 (2009).
- Müller, F., Zaucker, A. & Tora, L. Developmental regulation of transcription initiation: more than just changing the actors. *Curr. Opin. Genet. Dev.* **20**, 533–540 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank L. Cochella and O. Bell for comments on the manuscript. Deep sequencing was performed at the CSF Next-Generation Sequencing Unit (<http://csf.ac.at>). M.A.Z. was supported by the Austrian Science Fund (FWF, F4303-B09) and C.D.A., K.S., M.R. and O.F. by a European Research Council Starting Grant (no. 242922) awarded to A.S. Basic research at the Research Institute of Molecular Pathology is supported by Boehringer Ingelheim GmbH.

Author Contributions M.A.Z., C.D.A. and A.S. conceived the project. C.D.A., K.S., M.P., M.R. and O.F. performed the experiments and M.A.Z. the computational analyses. M.A.Z., C.D.A. and A.S. wrote the manuscript.

Author Information All deep sequencing data are available at <http://www.starklab.org> and have been deposited in the Gene Expression Omnibus database under accession numbers GSE40739 and GSE57876. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.S. (stark@starklab.org).

METHODS

hkCP STARR-seq vector. We derived the hkCP STARR-seq vector from the original STARR-seq vector¹² by replacing the DSCP sequence with the sequence of the *RpS12* core promoter (−50 to +50 bp relative to the TSS; TTGTACCAATAGCT AAAAAGTACATCTCCAGCGCCATGCCGATTTTGTCTCTTTCTTTCCG GTTGTCAAAAGGTACAGATGCTTGGATTTTATTCTC). The STARR-seq vectors are available subject to a material transfer agreement (MTA). For both STARR-seq vectors, we confirmed that transcription initiates from within the respective core promoters' Inr (DSCP) and TCT (*RpS12*) motifs by 5' rapid amplification of cDNA ends (RACE; Extended Data Fig. 2). All other STARR-seq vectors were derived from the hkCP STARR-seq vector by replacing the 100 bp sequence encompassing the *RpS12* core promoter by the sequences indicated in Supplementary Table 15 using the BglII and SbfI restriction sites.

hkCP and dCP luciferase vectors. For the dCP luciferase vector, the SV40 promoter of the pGL3-Promoter Vector (Promega) was replaced by the DSCP¹¹ and a Gateway cassette was inserted downstream of the luciferase gene and the SV40 polyA-signal into the AfeI restriction site, to allow Gateway LR cloning of candidate sequences¹². For the hkCP luciferase vector, the SV40 promoter and the sequence until the translation start codon of the luciferase gene was replaced by the sequence encompassing the TSS of *RpS12* from −50 bp until its translation start codon: TTGTACCAATAGCTAAAACTCACATCTCCAGCGCCATGCCGATTTTGTCTCTTTCTTTCCGGTTGTCAAAGGTACAGATGCTTGGATTTTATTTCTCGAAATGAAGAGGTTTCTTATCGAAAATGTAATAAATATGAACAATTAACATCTTTTCCAGTCAGTCATCCTTAACCGCAGAACA. Constructs are available subject to an MTA.

Intrinsic activity of core promoters. All core promoters used in this study were cloned into the dCP luciferase vector (without the Gateway cassette), replacing the DSCP between the BglII and SbfI restriction site with the respective core promoter. For each core promoter, the intrinsic (or basal) activity was measured as firefly luciferase activity and is presented as relative luciferase units, normalized to *Renilla* luciferase signals.

Genome-wide STARR-seq screens. STARR-seq enhancer screens using the core promoters of *RpS12* (hkCP), *NipB*, *x16*, and *eEF1δ* (Supplementary Table 15) were performed in two biological replicates (independent transfections) as described previously¹² with the following exceptions. First, 1.6×10^9 S2 cells and OSCs³¹ were transfected per biological replicate. Second, first-strand cDNA synthesis was performed in 30–60 reactions with the STARR-seq RT primer (CTCATCAATGTATCTTATCATGTCTG) as reverse transcription primer. Last, next-generation sequencing (NGS) was performed on an Illumina HiSeq 2000 machine using multiplexing according to the manufacturer's instructions. STARR-seq data using the DSCP (dCP STARR-seq) and *Hsp70* core promoters are from ref. 12, but were reanalysed using the same pipeline as for hkCP STARR-seq.

Focused STARR-seq BAC screens. The DSCP is a 137-nucleotide-long synthetic core promoter derived from the core promoter of *even-skipped* (*eve*)¹¹. To assess the functional similarity of the DSCP, its 137-nucleotide-long wild-type counterpart from the *eve* locus, and a version defined identically to all other core promoter used here (−50 to +50 nucleotides around the TSS), we performed STARR-seq screens with libraries derived from 29 different BACs containing a total of ~5 Mb of *D. melanogaster* genomic DNA (Supplementary Table 16). For comparison, we also screened all other core promoters with this library. For library cloning, all BACs were grown in individual bacterial cultures and were then mixed equally according to measurements of their optical density at 600 nm ($OD_{600\text{nm}}$) before BAC DNA isolation to achieve an equal distribution of all BACs. BAC DNA extraction, sonication and adaptor ligation was performed as described¹² and the same adaptor-ligated and PCR-amplified BAC DNA was used to clone all focused STARR-seq libraries. Per STARR-seq vector, four In-Fusion reactions were performed, which allowed five transformation reactions as described¹². Each library was grown in 4 l liquid culture (LB medium) to an $OD_{600\text{nm}}$ of 2.0–2.5. Each BAC library was screened as described earlier for the genome-wide screens; however, only 1×10^8 S2 cells were used, accounting for the less complex library. Similarly, the number of reactions for all subsequent steps of the STARR-seq protocol was reduced fourfold.

Luciferase reporter assays. Luciferase assays were performed as described previously¹² with the exception that the candidate enhancers were cloned downstream of the luciferase gene and the polyA signal, more than 2 kb away from the respective core promoter (*RpS12* or DSCP). Candidate enhancers were selected manually based on different criteria to allow the systematic assessment of several aspects of this study, including enhancers that were (1) specific to one of the two different core promoters (24 hkCP and 12 dCP enhancers) or found in both screens (7 shared enhancers); (2) located proximally (17) or distally (7) to the hkCP; and (3) of different strengths according to STARR-seq (ranks 18 to 1,044). We cloned all candidates as described¹² (for their genomic coordinates and primer sequences see Supplementary Table 17), picking initially one orientation towards the luciferase TSS randomly. However, to test the influence of TSSs contained in the candidate

sequences, we cloned and tested all TSS proximal candidates (hkCP_01 to hkCP_17) in both orientations using both core promoters. Candidate enhancers with DRE mutations were cloned from synthesized DNA fragments (GeneArt Strings; Supplementary Table 18). Candidates with DRE motifs that replace GAGA motifs were cloned similarly using synthesized DNA fragments (gBlocks) obtained from Integrated DNA Technologies (Supplementary Table 19). We also added an array of 6× DRE motifs into the AfeI restriction site of the dCP and hkCP luciferase vectors and cloned dCP_01 to dCP_11 into the middle of the DRE motif array (using AfeI) of the hkCP luciferase vector, such that these sequences were each flanked by three DRE motifs (Supplementary Table 19).

Luciferase assay data analysis. For all luciferase assays, we calculated standard deviations and one-sided Student's *t*-tests from three biological replicates (independent transfections). Core promoters have intrinsic (basal) activities that can differ between different core promoters. Therefore, when comparing enhancer activities for different core promoters, normalization to the core promoters' intrinsic activities is required, which we assessed with three different negative control fragments (nine biological replicates in total). For all measurements, we normalized firefly luciferase values first to *Renilla* luciferase values (controlling for transfection efficiency) and then to the normalized luciferase values of the three negative control sequences. Candidates with a significant ($P < 0.05$) enrichment greater than 1.5 fold over negative were considered positive.

5' RACE of STARR-seq transcripts. To determine the exact TSSs of hkCP and dCP within the STARR-seq vectors we performed 5' RACE of STARR-seq transcripts using one enhancer for each (an intergenic enhancer of *TpnC41C* for hkCP and an intronic enhancer of *zfh1* (shared_01 from ref. 12) for dCP) which we cloned with EcoRV at the position of the selection cassette used during library cloning (Supplementary Table 20). We transfected 3.2×10^7 cells with each of the constructs and isolated total RNA using the RNeasy mini prep kit (Qiagen; two columns per construct) followed by polyA+ RNA isolation using oligo-dT Dynabeads (Life Technologies) according to the manufacturer's instructions. We then performed 5' RACE for both samples using the FirstChoice RLM-RACE Kit (Ambion; catalogue no. AM1700) according to the manufacturer's instructions. To reflect RNA processing of the STARR-seq pipeline, reverse transcription was, however, performed using SuperscriptIII (Invitrogen) according to the manufacturer's instructions and using the reverse transcription primer GFP-RT (Supplementary Table 20) as a gene-specific primer (using RNA amounts according to the FirstChoice manual). The first PCR was performed with the manufacturer-provided 5' RACE Outer Primer and the transcript-specific primer RACE-01-rv, using 2× KAPA HiFi Hot Start Ready Mix (98 °C for 45 s; followed by 35 cycles of 98 °C for 15 s, 69 °C for 30 s, 72 °C for 30 s) with 1 µl of cDNA as template. The nested PCR was performed similarly (primer: 5' RACE Inner Primer and RACE-02-rv; 98 °C for 45 s; followed by 30 cycles of 98 °C for 15 s, 67 °C for 30 s, 72 °C for 10 s). The PCR products were visualized on a 1% agarose gel. The PCR products for both samples were Sanger sequenced using the primer GFP-seq-rv (for all primer sequences see Supplementary Table 20).

STARR-seq NGS data processing. Paired-end STARR-seq and input read processing was performed as described³². The NGS data for dCP (DSCP) and *Hsp70* were obtained from ref. 12 and reanalysed. In the same cell line, a hkCP peak is considered to be 'specific' if the 501 bp window centred at the peak summit does not overlap with any such window for dCP peaks, and vice versa (note that this is only applied within each cell type, such that comparisons across cell types are not influenced). For screens with the BAC-derived libraries, we considered only fragments that originated from the BACs used and determined the relative abundance of each BAC from the NGS data of the respective inputs only. On the basis of this, we then adjusted both inputs and STARR-seq NGS data such that all BACs were equally represented and analysed the data as described earlier.

Venn diagrams and peak intersection. We used the same intersection method as described earlier, and plotted the Venn diagrams with areas proportional to the number of peaks.

Scatter plots. We calculated the STARR-seq enrichment over input at the summit positions of both data sets that were to be compared, using a pseudo count of 1, and computed the log₂ of corrected ratio as described¹². This plots one data point for each enhancer—even for closely spaced ones—exactly at the enhancer's summit position. For visualizing replicates, we called peaks on the merged data sets and plotted the values from both replicates at these peaks' summits.

Enhancer-to-gene assignment. We performed three different strategies of enhancer-to-gene assignments: (1) 'closest TSS', whereby an enhancer is assigned to the closest TSS of an annotated transcript; (2) '1 kb TSS', whereby an enhancer is assigned to all TSSs that are within 1 kb; and (3) 'gene loci', whereby an enhancer is assigned to a gene provided that it falls within 5 kb upstream from the TSS, within the gene body itself, or 2 kb downstream of the gene (multiple assigned genes are possible). In all cases we used annotation from *D. melanogaster* FlyBase release 5.50.

Genomic distribution. We assigned a unique annotation for each nucleotide in the genome by using the following priority order: coding sequence (CDS), core promoter (± 50 bp around TSS), 5' UTR, 3' UTR, first intron, intron, proximal promoter (200 bp upstream of a TSS), intergenic region. We then assigned each peak to one of these categories by the annotation of the peak's summit.

GO analysis. We assessed whether genes assigned to hkCP or dCP enhancers were enriched for particular GO categories³³ by calculating hypergeometric *P* values for all categories, which we corrected for multiple comparisons (FDR-type correction in R). We then sorted all categories according to *P* values of overrepresentation, selected the top 100 of either hkCP or dCP, and removed redundant categories manually. For each category, we calculated $\log_{10}(P\text{-value underrepresentation}) - \log_{10}(P\text{-value overrepresentation})$, and sorted the terms in a descending order of difference between hkCP and dCP values. The colour intensity of the heat maps represents $\log_{10}(P\text{-value underrepresentation}) - \log_{10}(P\text{-value overrepresentation})$.

Gene expression analysis. We analysed enrichment in ubiquitous versus tissue-specific gene expression sets as described for the GO analysis above. To define the gene sets based on an *in situ* hybridization data set of fly embryos (BDGP³⁴), we first removed maternal (stages 1 to 3) annotations, as well as genes with the annotation 'no staining' in all stages. We required each gene to have annotations for at least three stage groupings. We called a gene 'tissue specific' if at most one of these annotations contains the word 'ubiquitous', and called it 'ubiquitous' if at least 60% of them contain word 'ubiquitous'. We also defined gene sets based on microarray data sets from dissected fly tissues (FlyAtlas³⁵). We defined genes as 'ubiquitous' if their expression does not change more than twofold compared with the whole fly for at least 15 out of 23 tissues. For this, we used the ratios and 'change_direction' calls from FlyAtlas directly and did not consider cell lines and carcasses. We similarly defined genes to be 'tissue specific' if they change more than twofold in at least three tissues. We do not consider genes with multiple conflicting entries as they can result from the use of multiple probes and removed genes that overlapped between the 'ubiquitous' and 'tissue-specific' gene sets from both sets.

Transcription factor motif and core promoter element enrichment analysis. We used previously employed position weight matrices (PWMs) for different transcription factors¹³ with a cut-off of $4^{-6} = 2.4 \times 10^{-4}$. We selected random control regions by controlling for genomic and chromosome distribution, and required that they did not overlap with any peak. We scored each motif for its enrichment in 401 bp windows centred on the peak summits by multiple testing (FDR) corrected hypergeometric *P* values. We considered only motifs that showed $\log_2(\text{confidence ratio of motif counts in peak windows/motif counts in random control regions}) > 1$ and *P* value < 0.01 in hkCP or dCP enhancers (or both) and reduced motif redundancy by removing highly similar motifs as in ref. 13 and references therein. We sorted the motifs in a descending order by difference in $\log_2(\text{hkCP enrichment}) - \log_2(\text{dCP enrichment})$. When assessing whether the observed motif distribution persisted for distal enhancers (Extended Data Fig. 10a), we kept the motifs and their order as in Fig. 5a and only re-evaluated their enrichment in distal enhancers. The colour intensity of the heat maps represents $\log_2(\text{confidence ratio of motif counts in peak windows/motif counts in random control regions})$. We used previously published PWMs or created PWMs from published nucleotide counts for TATA box, Inr, MTE, DPE and Ohler motifs¹⁶ 1, 5, 6, 7 and the TCT motif⁸ restricted to 8 bp. We scanned for motif occurrences using MAST from the MEME suite³⁶ (version 4.9.0) and parameters that ensured specificity and sensitivity for each motif (Supplementary Table 21). For enhancer-to-gene assignment methods 1 and 2 described earlier, we determined the presence of each core promoter element in the core promoter region of all genes uniquely assigned to either hkCP or dCP enhancers, respectively. For assignment method 3, we took the core promoter elements of the TSSs of the longest messenger RNA isoform. We assessed the differential distribution of each core promoter element between the core promoters assigned to hkCP or dCP enhancers by confidence ratios and hypergeometric *P* values.

Transcription factor motif and core promoter element *de novo* discovery. We used MEME³⁶ (version 4.9.0) to discover *de novo* motifs with lengths between 5 and 8 nucleotides in the enhancer regions we identified using STARR-seq and in

the core promoter regions around the nearest annotated transcription TSS. We provide all discovered motifs in Supplementary Table 22.

Core promoter similarity heat map. For all pairs of core promoters, we computed pair-wise PCCs between the respective STARR-seq fragment coverages at the summits of all peaks called in either of the two screens genome wide. We performed hierarchical clustering (complete linkage) in R, directly using the computed PCC values as similarities.

STARR-seq enrichment heat map. We computed the \log_2 of the corrected STARR-seq enrichment over input as described earlier, but for each nucleotide in a 20 kb window around all reference peak summit positions, and down-sampled the data points 50-fold by calculating one average data point per 50 nucleotides.

STARR-seq enrichment meta-profiles around TSSs. We calculated corrected STARR-seq enrichments (\log_2) as for the heat maps, but for 20 kb windows around TSSs, selected according to their core promoter motif content (see Extended Data Figs 4 and 8), corrected for the orientation of the TSSs within the genomic sequence. We then calculated the average for each position along the *x*-axis.

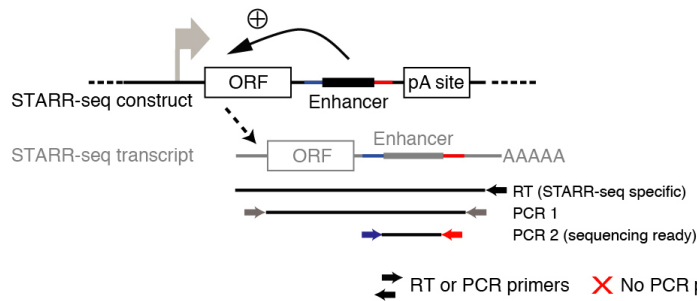
Boxplot. We obtained Dref ChIP-seq and input data (from Kc167 cells) from ref. 21 (Gene Expression Omnibus accession numbers GSM977024 and GSM762849) and mapped the 36-nucleotide reads using bowtie³⁷ (version 0.12.9) with the following parameters: -p 4 -q -v 3 -m 1 --best --strata --quiet. We extended the reads to 150 bp, calculated the coverage for ChIP-seq and input at the STARR-seq peak summit, normalized the value to the number of input fragments, added a pseudo count of 1, and computed the confidence ratio of ChIP-seq over input. For the Trl ChIP-chip data obtained from ref. 22, we used the signal of the chip-array probe at the peak summit if available or inferred the signal by linear extrapolation from the two nearest flanking probes (one on each side) provided that they were both within 10 nucleotides of the peak summit. We calculated statistical significance via Wilcoxon's paired rank tests.

Coordinate intersections. We performed genomic coordinate intersections using the BEDTools suite³⁸ (version 2.17.0).

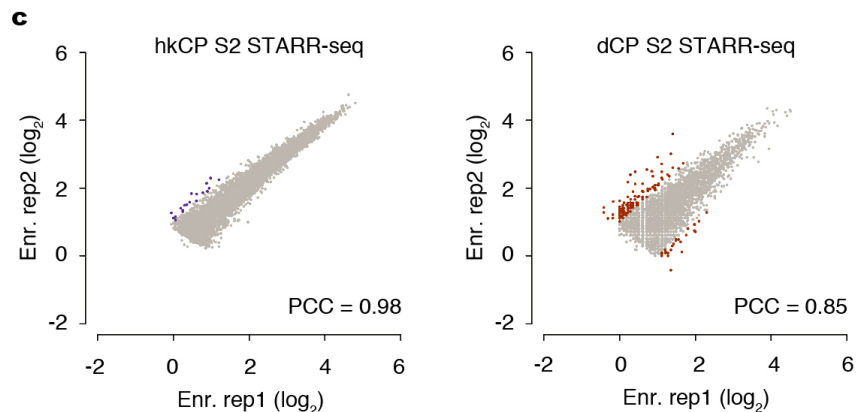
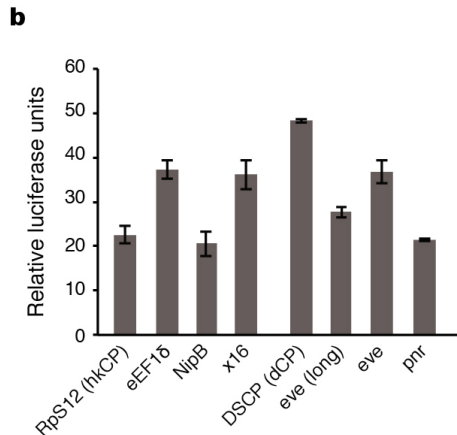
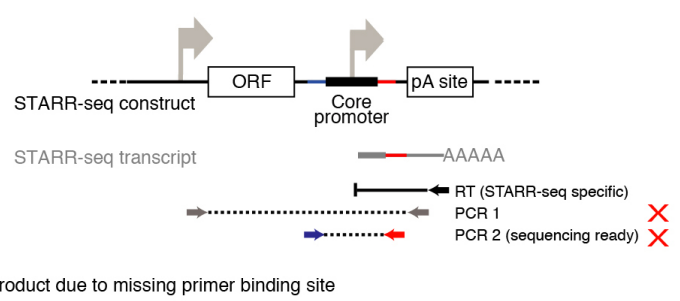
Statistics. We performed all statistical calculations and created graphical displays with R³⁹.

- Saito, K. *et al.* A regulatory circuit for *piwi* by the large Maf gene *traffic jam* in *Drosophila*. *Nature* **461**, 1296–1299 (2009).
- Arnold, C. D. *et al.* Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nature Genet.* **46**, 685–692 (2014).
- Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature Genet.* **25**, 25–29 (2000).
- Tomanek, P. *et al.* Global analysis of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol.* **8**, R145 (2007).
- Chintapalli, V. R., Wang, J. & Dow, J. A. T. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nature Genet.* **39**, 715–720 (2007).
- Bailey, T. L. & Gribskov, M. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics* **14**, 48–54 (1998).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2010).
- Zeitlinger, J. & Stark, A. Developmental gene regulation in the era of genomics. *Dev. Biol.* **339**, 230–239 (2010).
- Kvon, E. Z. *et al.* Genome-scale functional characterization of *Drosophila* developmental enhancers *in vivo*. *Nature* **512**, 91–95 (2014).
- Soler, E. *et al.* The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev.* **24**, 277–289 (2010).
- Chen, K. *et al.* A global change in RNA polymerase II pausing during the *Drosophila* midblastula transition. *eLife* **2**, e00861 (2013).
- Lagha, M. *et al.* Paused Pol II coordinates tissue morphogenesis in the *Drosophila* embryo. *Cell* **153**, 976–987 (2013).
- Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).

a Candidate fragment is an enhancer → detected



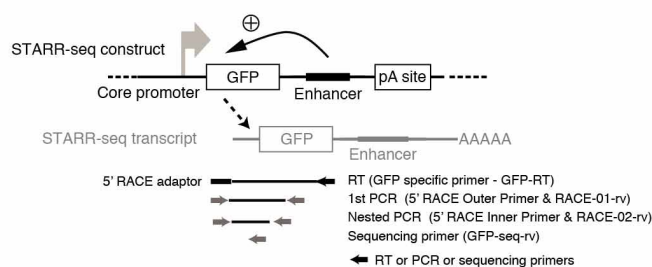
Candidate fragment is a core promoter → not detected



Extended Data Figure 1 | Set-up of STARR-seq with different core promoters.

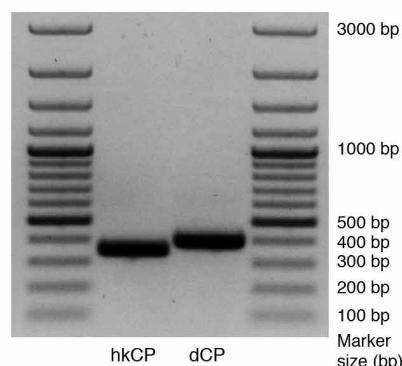
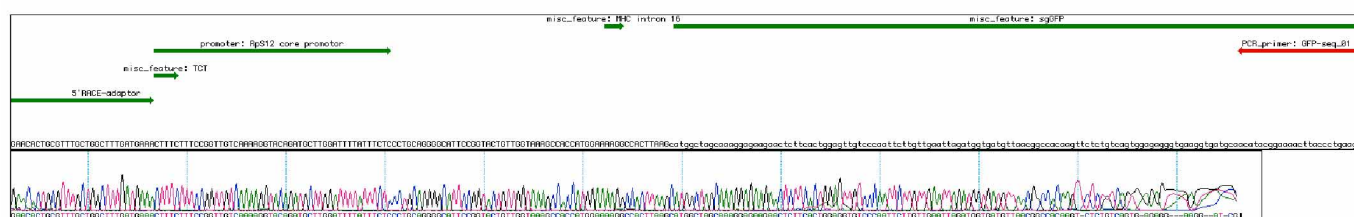
a, STARR-seq detects enhancers but no promoters (reproduced with permission from ref. 12). Left, STARR-seq couples the enhancer activities of candidate fragments to the sequences of the candidates *in cis* by placing the candidates to a position within the reporter transcript. Enhancer activities can therefore be assessed by the presence of candidates among cellular messenger RNAs, which allows the parallel assessment of millions of candidates, enabling genome-wide screens. Sequences that activate transcription from the intended core promoter of the STARR-seq vector lead to a full-length reporter transcript and can be detected by STARR-seq. Shown are the reverse transcription (RT) and nested polymerase chain reaction (PCR) steps of the STARR-seq reporter RNA processing protocol that ensure this. Right, in contrast, STARR-seq does not detect truncated transcripts that result if a candidate fragment functions as a promoter to initiate transcription. Thus, core-promoter-containing (that is, TSS-overlapping) sequences that are detected by STARR-seq exhibit enhancer activity as they can activate transcription from a remote position, in addition to their ability to serve as core

promoters endogenously¹². **b**, Luciferase signals (firefly/*Renilla*) assessing the intrinsic (or basal) activity of the core promoters used in this study. The luciferase reporter constructs do not contain any enhancer and differ only in the respective core promoter sequences. The basal activities differ as expected, but do not differ consistently between housekeeping (*Rps12*, *eEF1δ*, *NipB*, *x16*) and developmental (*DSCP*, *eve* (long), *eve* and *pnr*) core promoters, nor between core promoters for which the STARR-seq screens appear most similar (for example, *Rps12* and *eEF1δ*; see Fig. 3). Note that all luciferase assays and STARR-seq screens are corrected for differences in intrinsic activity. **c**, Reproducibility of hkCP and dCP STARR-seq in *D. melanogaster* S2 cells. The reproducibility of hkCP and dCP STARR-seq as assessed by the STARR-seq enrichments (replicate 1 versus 2) at the summits of enhancer peaks called in the merged experiments (hkCP: 5,956; dCP: 5,408). Scatter plots are enlarged versions of the insets in Fig. 1d. “Enr. rep X”, STARR-seq enrichment in replicate X. Note that the raw data for dCP have been re-analysed from ref. 12.

a Schematic overview of 5' RACE on the STARR-seq vectors

Core-promoter–enhancer pairs used:

hkCP (*RpS12* core promoter) and hkCP_19 (expected size of nested PCR product: ~350bp)
dCP (DSCP) and intronic enhancer of *zfh1* (expected size of nested PCR product: ~400bp)

b Agarose gel electrophoresis of 5' RACE nested PCRs (PCR2)**c** Sequence alignment and chromatogram of Sanger sequencing of the PCR product (nested PCR) of 5' RACE of the hkCP STARR-seq vector

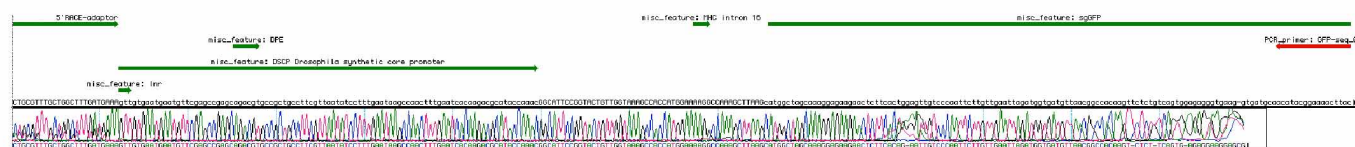
RpS12 core promoter (STARR-seq vector)

Full sequence (-50 to +50)
Transcribed sequence

ttgtaccaaatagctaaaaactcacatctccagcgccatgcgattttgTCTCTTTTctttccggtgtgcaaaaggtacagatgcttgattttattttc

TCT

TSS on plasmid

d Sequence alignment and chromatogram of Sanger sequencing of the PCR product (nested PCR) of 5' RACE of the dCP STARR-seq vector

DSCP (STARR-seq vector)

TATA box

gagctcgccggggatcgagcgagcggtTATAAagggcggggtggctgagagcaTCAGTgtgaatgaatgtTCGAGCCGAGCAGAGCTgcccgtgccttcgttaataatctttgaataagcaactttgaatcacaagcgataccaaac

Inr

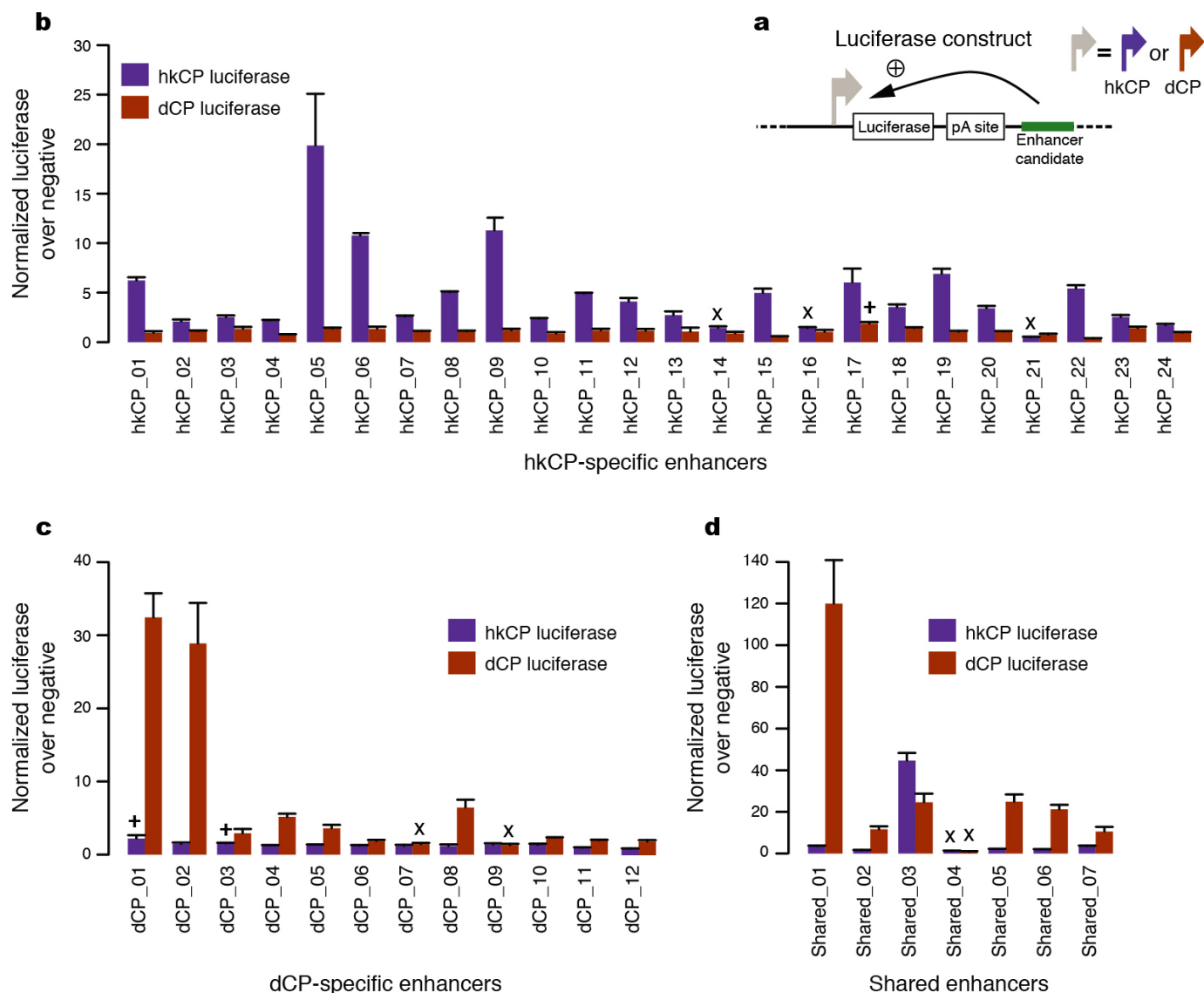
MTE DPE

Full sequence (-58 to +96)

Transcribed sequence

Extended Data Figure 2 | Transcription initiates within the core promoter of the STARR-seq construct. **a–d**, 5' Rapid amplification of cDNA ends (5' RACE) demonstrates that transcription initiates at the TCT and Inr motifs within the hkCP and dCP, respectively. **a**, Set-up of the 5' RACE experiment, including the STARR-seq plasmid, used here with two defined enhancers, the STARR-seq transcript and the location of all primers used to specifically amplify 5'-capped STARR-seq transcripts. **b**, 5' RACE nested PCR products separated on a 1% agarose gel. **c**, Screenshot of Sanger sequencing results

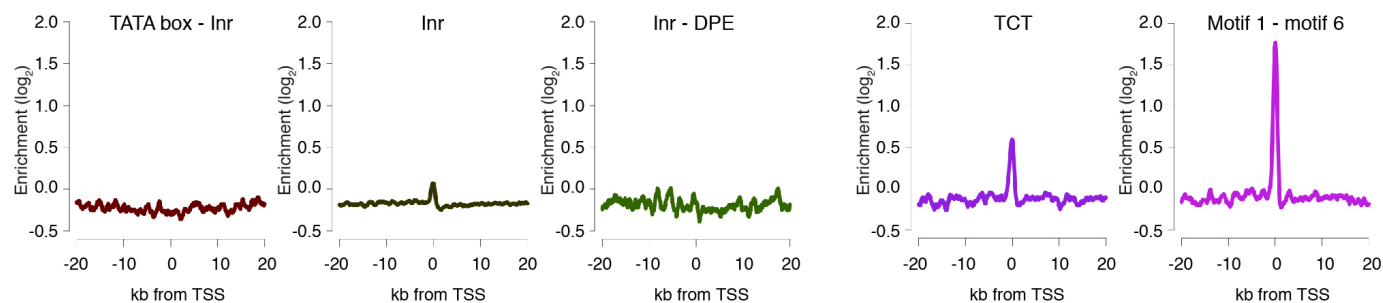
(chromatogram and called bases) compared with the template sequence. Annotations are shown in green, in the following order: 5' RACE adaptor, hkCP with TCT motif (only the part downstream of the TSS is annotated, as the 5' part is not present in the sequenced complementary DNA), spliced intron, green fluorescent protein (GFP); the sequencing primer is shown in red (top). Also shown is a version that displays the template and Sanger sequencing results for the core promoter region only (zoom in). **d**, Same as in **c** but for the dCP for which transcription initiates within the Inr motif.



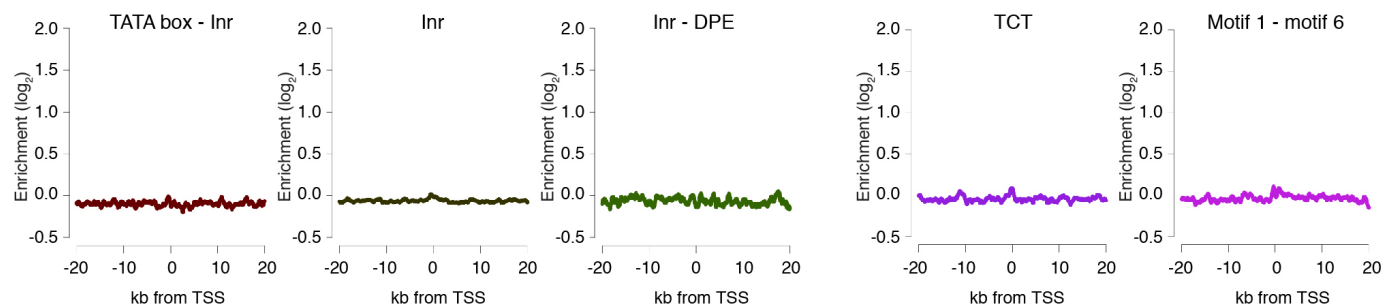
Extended Data Figure 3 | Specificity of hkCP and dCP enhancers to the hkCP and dCP assessed by luciferase assays. **a**, Luciferase reporter set-up with the hkCP or dCP (see also Fig. 1e). **b**, Luciferase signals of 24 hkCP-specific enhancers tested in a hkCP- (purple bars) as well as in a dCP-containing (brown bars) luciferase reporter. Twenty-one out of 24 hkCP enhancers showed luciferase activity (>1.5 fold over negative, $P < 0.05$ via one-sided unpaired Student's *t*-test, $n = 3$) with the hkCP, while only 1 out of 24 showed activity with the dCP (error bars are s.d. of three biological replicates, 'X' indicates candidates that are not active with the correct core promoter, and '+'

indicates candidates for which the activity with the wrong core promoter is above the threshold (note that the activity with the correct core promoter is still higher in all three cases). **c**, As in **b** but testing dCP-specific enhancers. Ten out of 12 are positive with the dCP whereas only 2 out of 12 are positive with the hkCP. **d**, As in **b** and **c** but testing shared enhancers that were found by STARR-seq with hkCP and dCP; 6 out of 7 are active with both core promoters. See Supplementary Table 17 for the genomic coordinates of the enhancers and the primers used to amplify them.

hkCP S2 STARR-seq signal around core promoter types

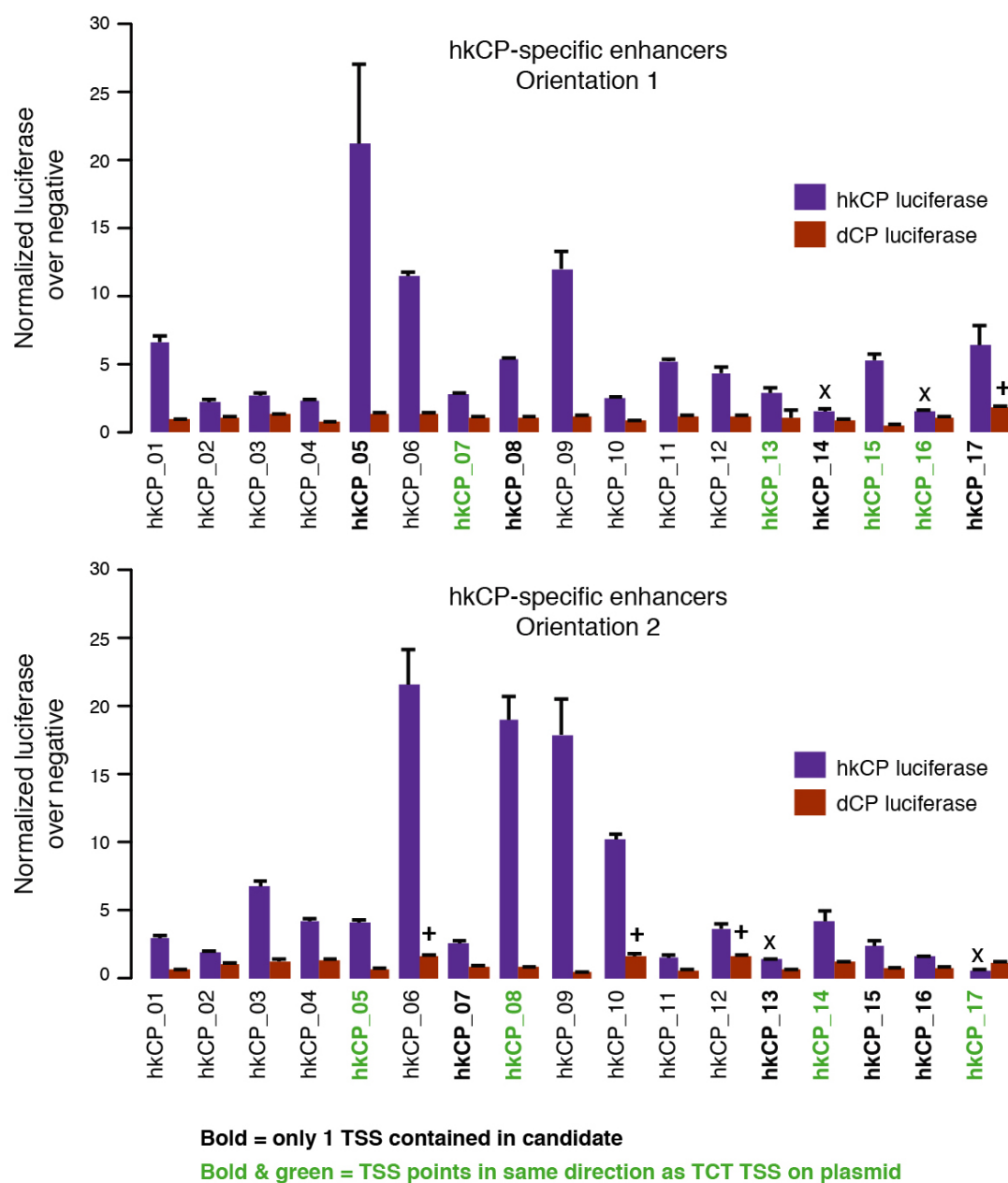


dCP S2 STARR-seq signal around core promoter types



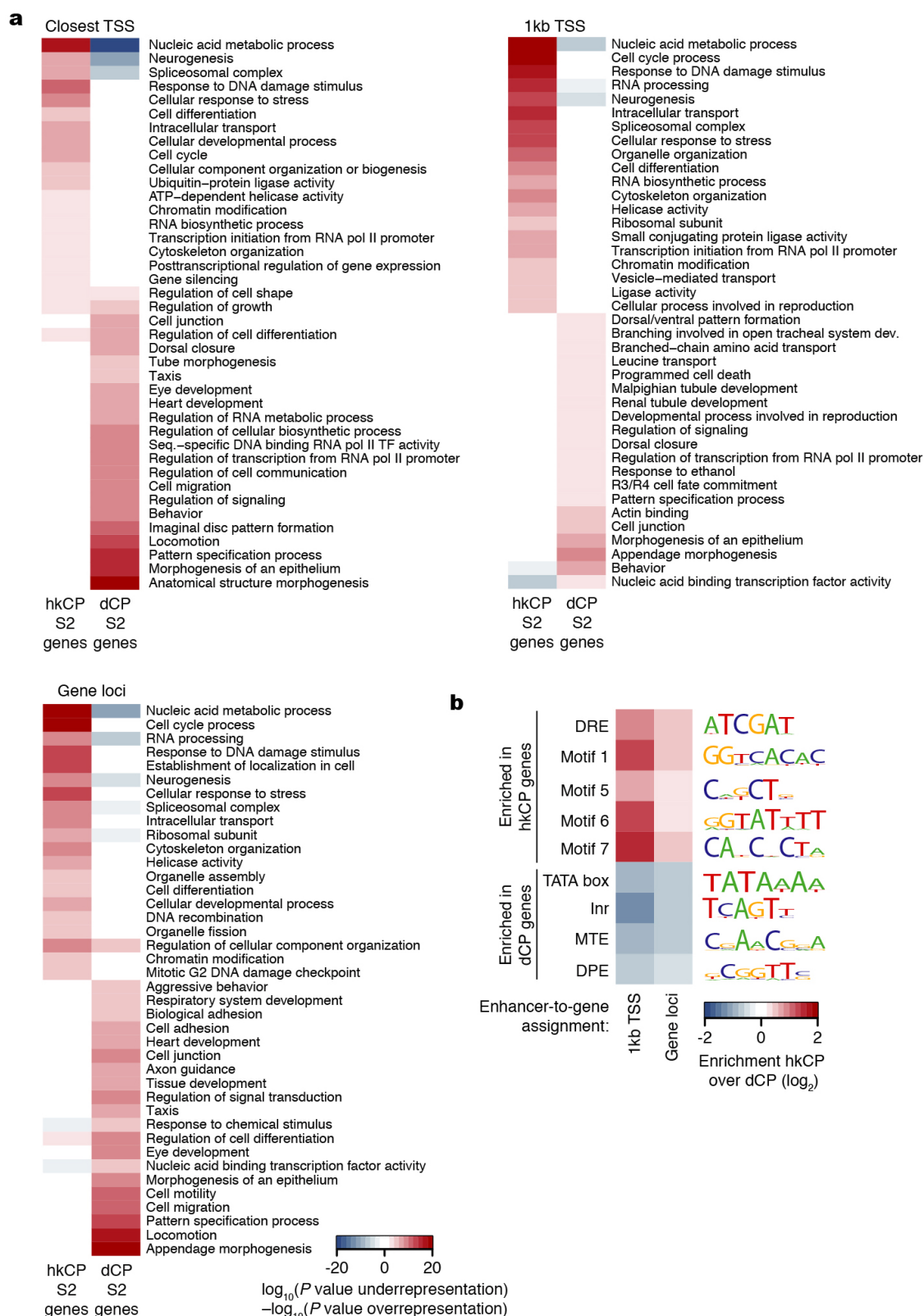
Extended Data Figure 4 | hkCP and dCP STARR-seq signal in S2 cells around different core promoter types. Average hkCP (top) and dCP (bottom) S2 STARR-seq enrichment in 40 kb intervals around TSSs that contain different combinations of known core promoter motifs. Shown are (left to right) TATA box–Inr (179 TSSs), Inr (that do not contain either TATA box or DPE; 1,901), Inr–DPE (100), TCT (303) and motif 1–motif 6 (266). According to

their motif contents, the first three are developmental-type core promoters and the last two are housekeeping-type core promoters. Indeed, only the housekeeping-type core promoters show a strong enrichment of hkCP S2 STARR-seq signals at the TSS, which is not seen for the dCP STARR-seq signal (owing to enhancer–core-promoter specificity) nor for the developmental-type core promoters (owing to the dCP enhancers location at more distal sites).



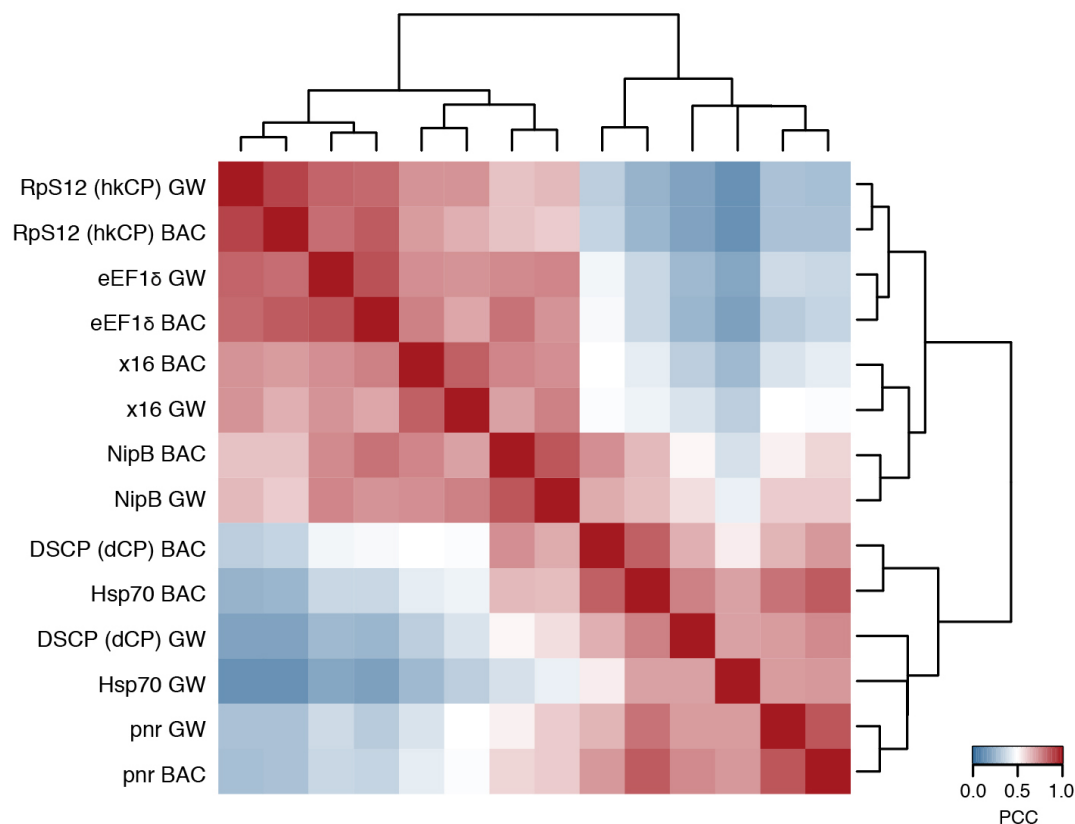
Extended Data Figure 5 | TSS-overlapping hkCP enhancers function independent of their orientation. Luciferase signals for all 17 TSS-overlapping hkCP enhancers (that is, containing one TSS or two divergent TSSs; see Supplementary Table 17) from Extended Data Fig. 3 cloned in the second orientation with respect to the TSS of the luciferase gene (bottom bar plot; the top bar plot corresponds to the initial orientation as in Extended Data Fig. 3 and is shown for comparison). In both orientations, 15 out of 17 enhancers showed activity towards the hkCP (details as in Extended Data Fig. 3). These results together with the findings in Extended Data Fig. 3 challenge the widespread notion that TSS-proximal sequences are promoters and even the concept of promoters more generally: sequences that

autonomously activate gene expression—and are therefore often termed promoters—might in fact be the combination of a core promoter and a proximal enhancer. The TSS-proximal location of many housekeeping enhancers might be evolutionarily more ancient, consistent with regulatory mechanisms in simple eukaryotes such as yeast. In contrast, enhancers of genes with more complex regulation are typically located more distally, potentially simply because the several different cell-type-specific enhancers of these genes would not all fit to positions near TSSs. Consistently, such genes frequently have larger intergenic and intragenic regions⁴⁰ known to accommodate enhancers with diverse activity patterns⁴¹.



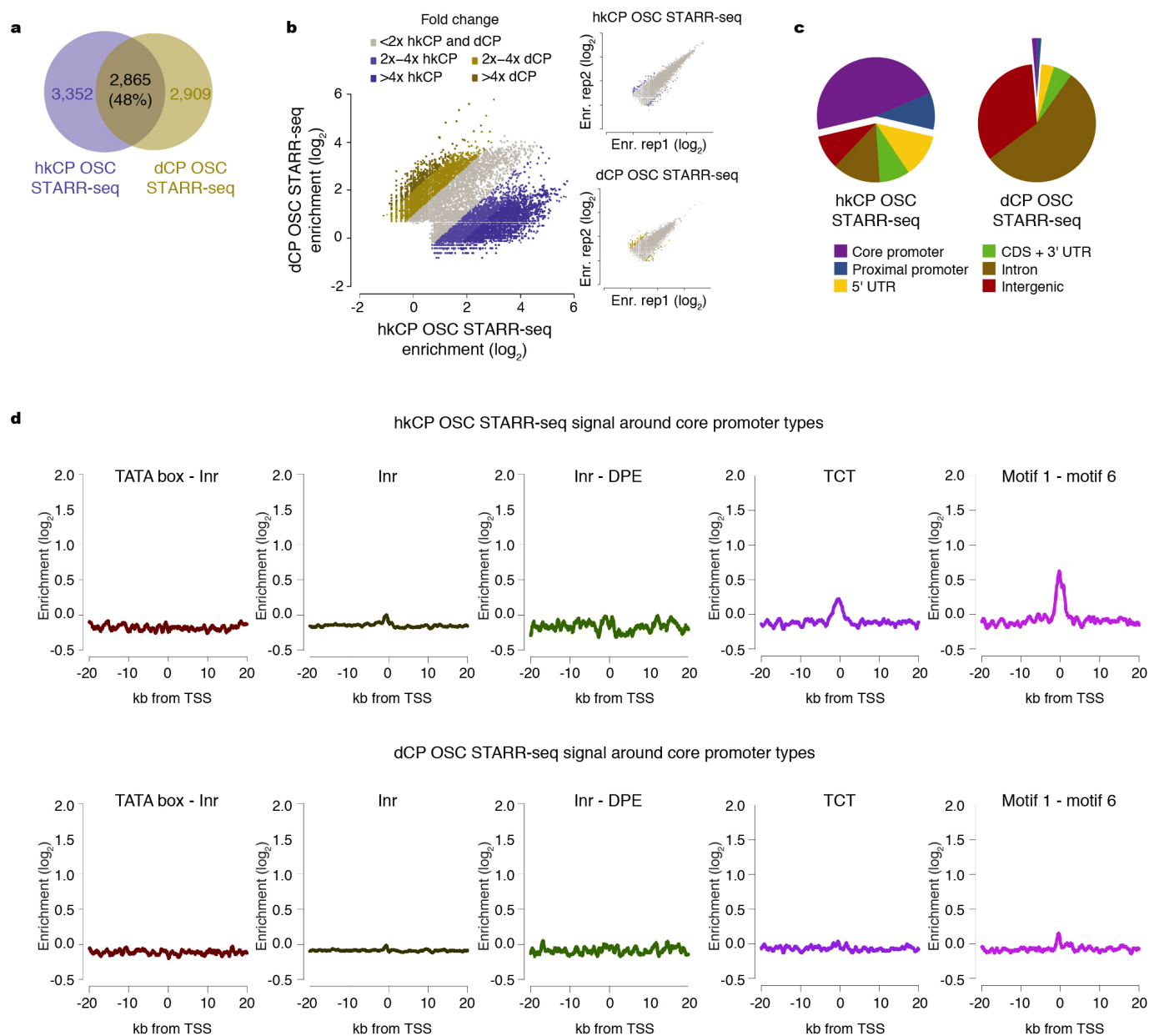
Extended Data Figure 6 | hkCP and dCP enhancers in S2 cells are associated with genes of different functions and core promoter elements. **a**, GO analysis of genes next to hkCP- and dCP-specific enhancers in S2 cells using different enhancer-to-gene assignment strategies (top left, 'closest TSS' as in Fig. 2; top right, '1 kb TSS'; bottom left, 'gene loci'; see Methods for details). Shown are 20 non-redundant GO categories selected from the 100 most significantly enriched categories associated with each enhancer class (see Supplementary

Tables 2–4 for all categories). **b**, Enrichment of core promoter elements at genes next to hkCP- and dCP-specific enhancers in S2 cells. Similar analysis as in Fig. 2e, but using different enhancer-to-gene assignment strategies (see Methods for details). Consistent with Fig. 2e, core promoters of genes assigned to hkCP-specific enhancers are enriched in motifs 1, 5, 6, 7 and DRE, while core promoters of genes assigned to dCP-specific enhancers are enriched for TATA box, Inr, MTE and DPE motifs, irrespective of the assignment strategy.



Extended Data Figure 7 | Housekeeping and developmental core promoters differ characteristically in their global enhancer preferences. As in Fig. 3b but including biological replicates with independently cloned focused bacterial artificial chromosome (BAC) libraries covering around 5 Mb of genomic sequence (BAC) and assessing the PCC at each position along these regions. GW, genome-wide screens as in Fig. 3b. The similarity observed for the TATA

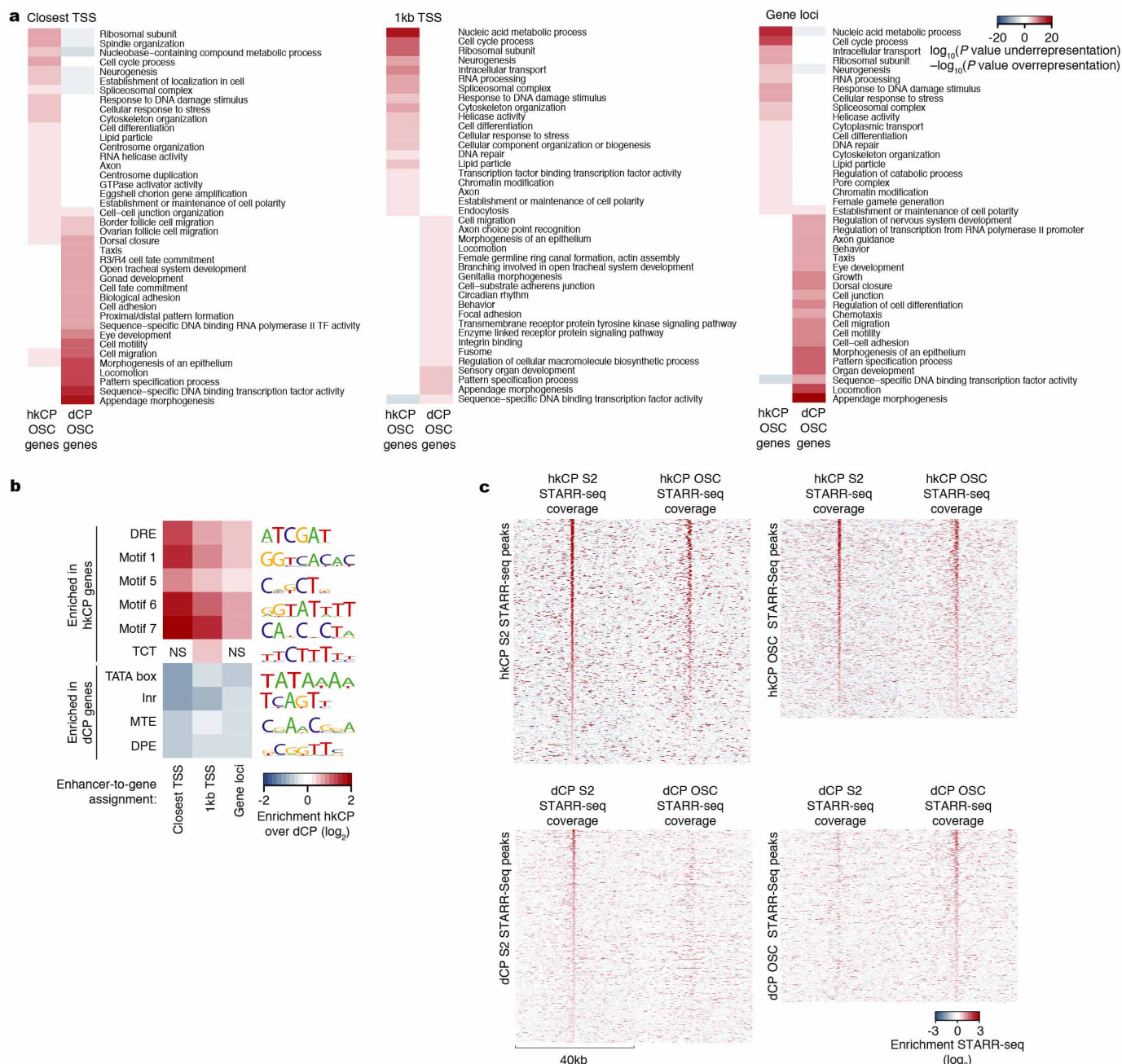
box- and DPE-containing core promoters (*Hsp70*, *pnr* and DSCP (dCP)) suggest that differences related to these core promoter elements might be more subtle or related to alternative mechanisms, including the potential preferences of more proximal or distal enhancers⁴² or RNA polymerase II pausing and the dynamics versus stochasticity of initiation and elongation^{43,44,45}.



Extended Data Figure 8 | hkCP and dCP enhancers differ in OSCs.

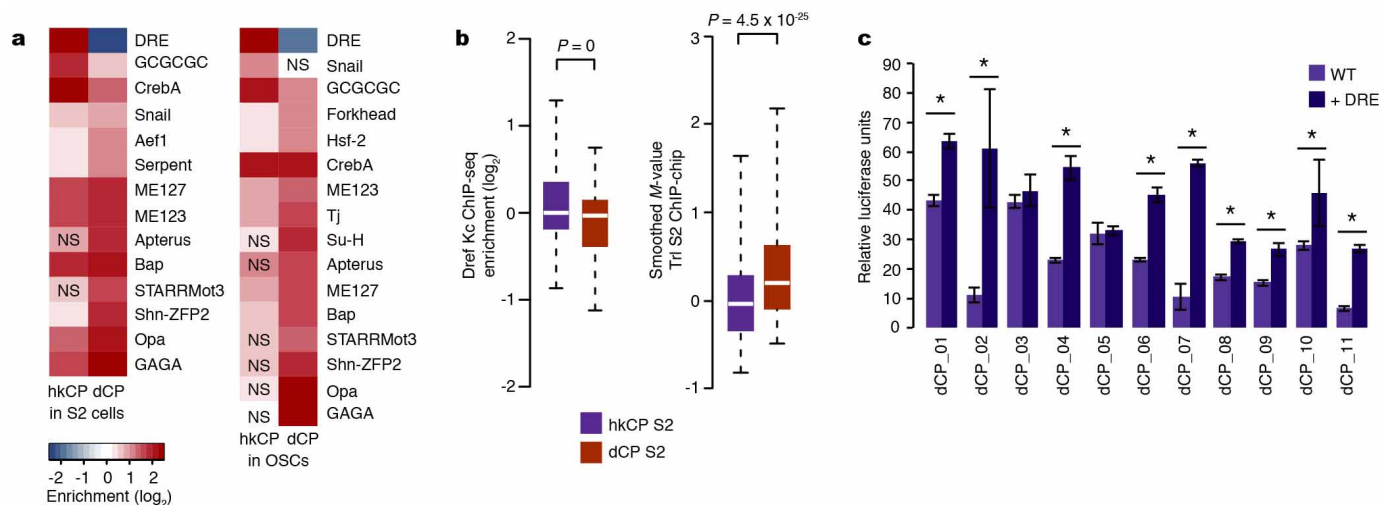
a, b, Different enhancers activate transcription from hkCP and dCP in OSCs. As Fig. 1c, d but for OSCs rather than S2 cells (data in bottom inset of **b** are re-analysed from ref. 12). **c**, Genomic distribution of hkCP and dCP

enhancers in OSCs. As Fig. 2a but for OSCs rather than S2 cells. **d**, hkCP and dCP STARR-seq signal in OSCs around different core promoter types. As Extended Data Fig. 4 but for OSCs rather than S2 cells.



Extended Data Figure 9 | Differences between hkCP and dCP enhancers in OSCs. **a**, GO analysis of genes next to hkCP- and dCP-specific enhancers in OSCs. As Extended Data Fig. 6a but for OSCs rather than S2 cells (see Supplementary Tables 8–10 for all categories). **b**, Enrichment of core promoter elements at genes next to hkCP- and dCP-specific enhancers in OSCs. As Fig. 2e

and Extended Data Fig. 6b but for OSCs rather than S2 cells. NS, not significant (hypergeometric $P > 0.05$). **c**, Heat maps of hkCP (top) and dCP (bottom) STARR-seq enrichments in S2 cells and OSCs. Heat maps on the left and right are centred on the summits of core-promoter-type-specific enhancers in S2 and OSCs, respectively.



Extended Data Figure 10 | The activities of hkCP and dCP enhancers are dependent on DRE and GAGA motifs, respectively. **a**, Differential motif enrichment in distally located hkCP- and dCP-specific enhancers (as in Fig. 5a but assessing enrichments of the same motif PWMs exclusively at distal enhancers >500 bp away from the closest TSSs). Key motifs including DRE and GAGA are also differentially enriched in distal hkCP- and dCP-specific enhancers. NS, not significant (FDR-corrected hypergeometric $P > 0.01$). S2 cells: hkCP $n = 790$, dCP $n = 3,013$; OSCs: hkCP $n = 556$, dCP $n = 2,555$. **b**, Distal hkCP- and dCP-specific enhancers are differentially bound by Dref and Trl, respectively. ChIP enrichments of Dref (left) and Trl (right) at S2 hkCP- and dCP-specific enhancers that are distal (>500 bp) from the closest TSSs. Equivalent to Fig. 5b, but considering exclusively TSS-distal enhancers to

exclude potentially confounding effects for TSS-proximal enhancers for which it is not possible to discern whether binding occurs due to the enhancer sequence or core promoter function. The differential binding between Dref and Trl to hkCP- and dCP-specific enhancers, respectively, is also found in Kc167 cells, in which the Dref ChIP-seq experiment had been performed (data not shown). **c**, Addition of DRE motifs to dCP enhancers increases their activity towards hkCP. Relative luciferase activity values (firefly/*Renilla*) for 11 dCP enhancers without DRE motifs (wild type (WT), light purple) and with 3 DRE motifs flanking the enhancers on each side (+DRE, dark purple). * $P < 0.05$, one-sided unpaired Student's *t*-test; error bars denote the s.d. of three biological replicates.

N^6 -methyladenosine-dependent RNA structural switches regulate RNA–protein interactions

Nian Liu¹, Qing Dai¹, Guanqun Zheng², Chuan He^{1,2,3,4}, Marc Parisien^{2†} & Tao Pan^{2,3}

RNA-binding proteins control many aspects of cellular biology through binding single-stranded RNA binding motifs (RBMs)^{1–3}. However, RBMs can be buried within their local RNA structures^{4–7}, thus inhibiting RNA–protein interactions. N^6 -methyladenosine (m^6A), the most abundant and dynamic internal modification in eukaryotic messenger RNA^{8–10}, can be selectively recognized by the YTHDF2 protein to affect the stability of cytoplasmic mRNAs¹⁵, but how m^6A achieves its wide-ranging physiological role needs further exploration. Here we show in human cells that m^6A controls the RNA-structure-dependent accessibility of RBMs to affect RNA–protein interactions for biological regulation; we term this mechanism ‘the m^6A -switch’. We found that m^6A alters the local structure in mRNA and long non-coding RNA (lncRNA) to facilitate binding of heterogeneous nuclear ribonucleoprotein C (HNRNPC), an abundant nuclear RNA-binding protein responsible for pre-mRNA processing^{20–24}. Combining

photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) and anti- m^6A immunoprecipitation (MeRIP) approaches enabled us to identify 39,060 m^6A -switches among HNRNPC-binding sites; and global m^6A reduction decreased HNRNPC binding at 2,798 high-confidence m^6A -switches. We determined that these m^6A -switch-regulated HNRNPC-binding activities affect the abundance as well as alternative splicing of target mRNAs, demonstrating the regulatory role of m^6A -switches on gene expression and RNA maturation. Our results illustrate how RNA-binding proteins gain regulated access to their RBMs through m^6A -dependent RNA structural remodelling, and provide a new direction for investigating RNA-modification-coded cellular biology.

Post-transcriptional m^6A RNA modification is indispensable for cell viability and development, yet its functional mechanisms are still poorly understood^{8–19}. We recently identified one m^6A site in a hairpin-stem

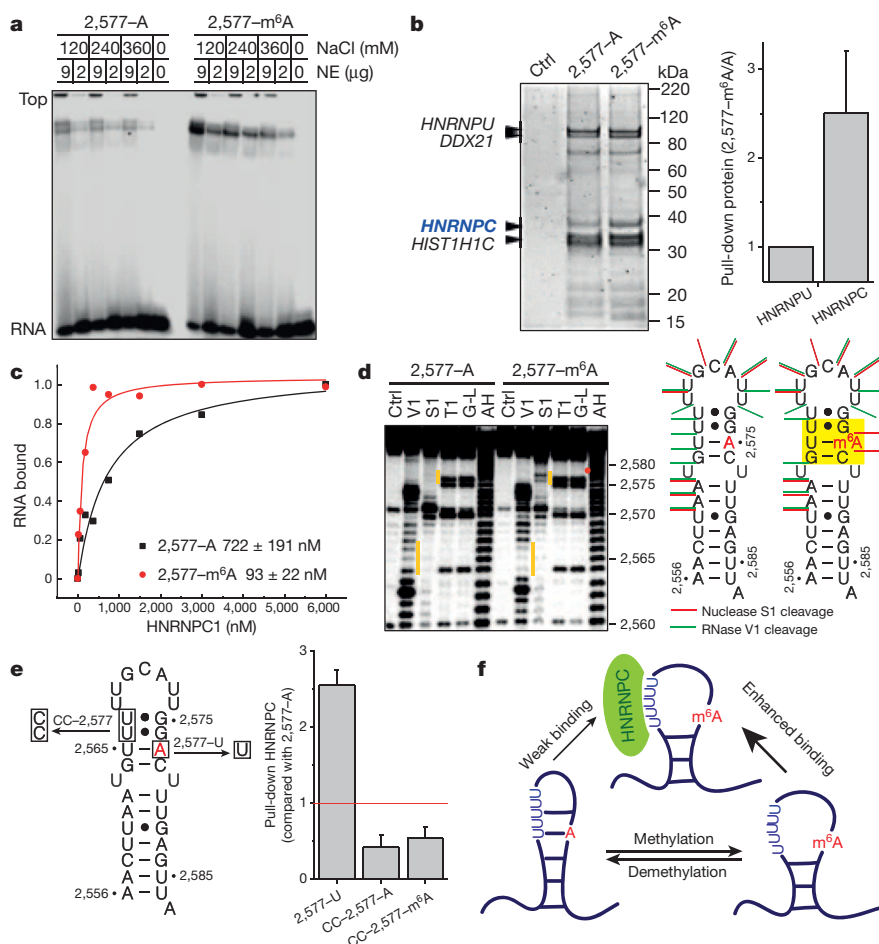


Figure 1 | m^6A alters RNA structure to enhance HNRNPC binding. **a**, m^6A increases binding with nuclear proteins. NE, nuclear extract. **b**, RNA pull down showing that HNRNPC preferably binds methylated RNA. Data are mean \pm standard deviation (s.d.); $n = 4$, biological replicates. Ctrl, control. **c**, Filter binding showing that m^6A increases HNRNPC1 binding with respective apparent dissociation constants (K_d) indicated at bottom right. Data are mean \pm s.d.; $n = 3$, technical replicates. **d**, RNA structural probing showing that m^6A disrupts local RNA structure, highlighted in yellow. AH, alkaline hydrolysis; Ctrl, no nuclease added; G-I, G-ladder; S1, nuclease S1 digestion; T1, RNase T1 digestion; V1, RNase V1 digestion. The orange bars mark the structurally altered RNA regions in the presence of m^6A (red dot). **e**, RNA pull down with mutated oligonucleotides. Data are mean \pm s.d.; $n = 3$, technical replicates. **f**, Illustration of the m^6A -switch model.

¹Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, USA. ²Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois 60637, USA. ³Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, USA. ⁴Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637, USA. [†]Present address: Faculty of Dentistry, McGill University, Montreal, Quebec H3A 0G1, Canada.

on the human lncRNA metastasis-associated lung adenocarcinoma transcript (*MALAT1*)²⁵ (Extended Data Fig. 1a). A native gel-shift assay indicated that this m⁶A residue increases the interaction of the RNA hairpin with proteins in the HeLa nuclear extract (Fig. 1a). RNA pull-down assays identified HNRNPC as the protein component of the nuclear extract that binds more strongly with the m⁶A-modified hairpin than the unmodified hairpin (Fig. 1b and Extended Data Fig. 1b, c). The m⁶A-enhanced interaction with the hairpins was validated qualitatively by ultraviolet crosslinking and quantitatively (~8-fold increase) by filter binding using recombinant HNRNPC1 protein (Fig. 1c and Extended Data Fig. 1d).

The HNRNPC protein belongs to the large family of ubiquitously expressed heterogeneous nuclear ribonucleoproteins that bind nascent RNA transcripts to affect pre-mRNA stability, splicing, export and translation^{20–24}. HNRNPC preferably binds single-stranded U-tracts (five or more contiguous uridines)^{20,23,24,26,27}. In the *MALAT1* hairpin, HNRNPC binds a U₅-tract that is half buried in the hairpin-stem opposing the 2,577-A/m⁶A site (Extended Data Fig. 1a, e).

Since m⁶A residues within RNA stems can destabilize the thermostability of model RNA duplexes²⁸, we hypothesized that the 2,577-m⁶A residue destabilizes this *MALAT1* hairpin-stem to make its opposing U-tract more single-stranded or accessible, thus enhancing its interaction with HNRNPC. We performed several experiments to validate

this hypothesis. First, according to the RNA structural probing assays, the m⁶A-modified hairpin showed significantly increased nuclease S1 digestion (single-strand specific) at the GAC (A = m⁶A) motif, as well as markedly decreased RNase V1 digestion (double-strand/stacking specific) at the U-tract opposing the GAC motif (Fig. 1d). The m⁶A residue markedly destabilized the stacking properties of the region centred around the U residue that pairs with 2,577-A/m⁶A (Extended Data Fig. 1f, g), which was also supported by the increased reactivity between CMCT and the U-tract bases in the presence of m⁶A (Extended Data Fig. 1h). Second, the 2,577-A-to-U mutation increased the HNRNPC pull-down amount from the nuclear extract, whereas U-to-C mutations in the U-tract significantly reduced the HNRNPC pull-down amount regardless of m⁶A modification (Fig. 1e). Third, the 2,577-A-to-U mutation increased the accessibility of the U-tract and enhanced HNRNPC binding by ~4-fold (Extended Data Fig. 2a–c). Binding results with four other mutated A/m⁶A oligonucleotides also supported the U-tract, with increased accessibility alone being sufficient to enhance HNRNPC binding (Extended Data Fig. 2d). Fourth, RNA terminal truncation followed by HNRNPC binding identified two pairs of truncated hairpins with highly accessible U-tracts, which improved HNRNPC binding significantly but independent of the m⁶A modification (Extended Data Fig. 2e–i). All these results confirmed that m⁶A modification can alter its local RNA structure and enhance the accessibility of its base-paired

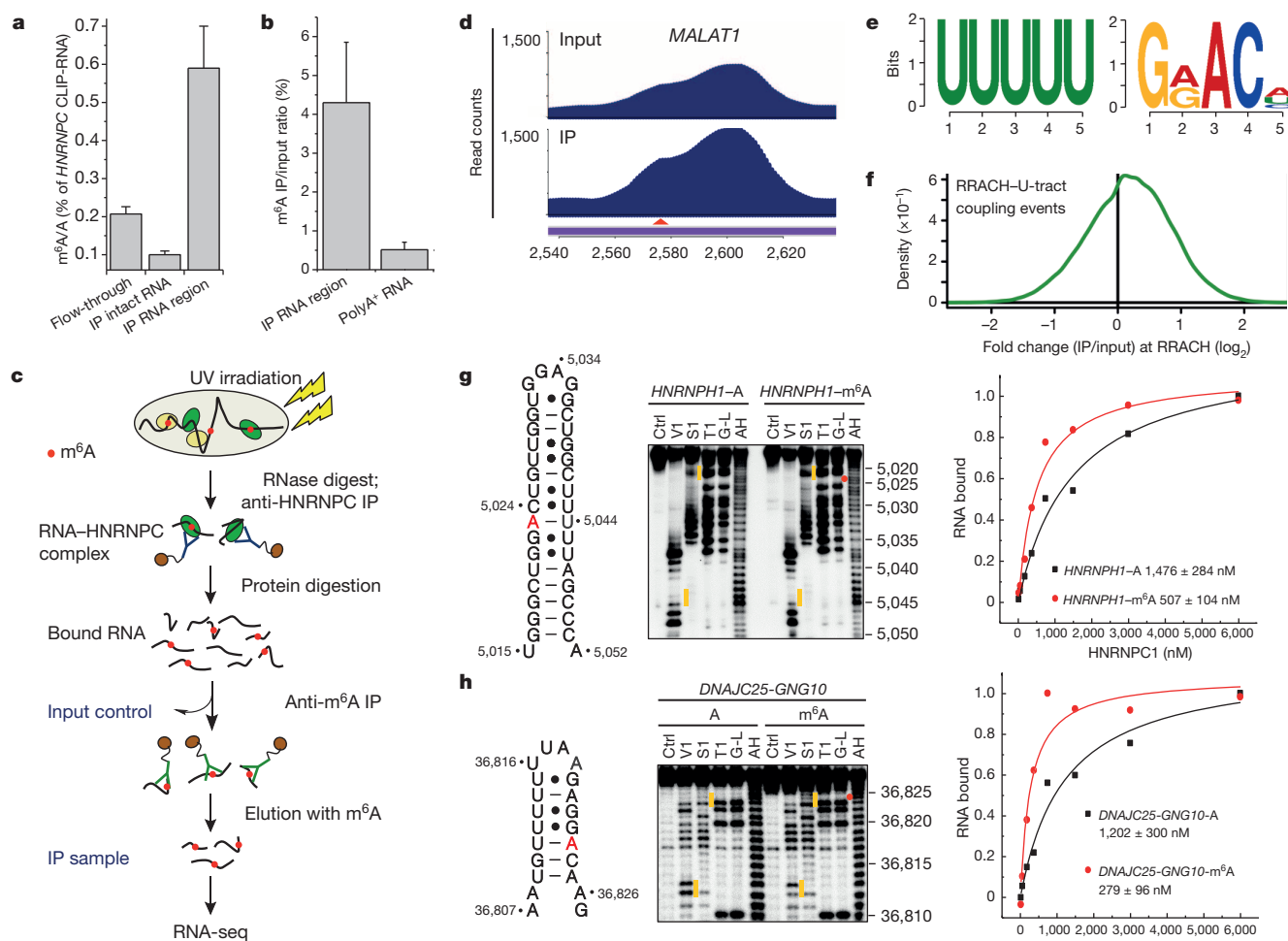


Figure 2 | PAR-CLIP-MeRIP identifies m⁶A-switches transcriptome wide. **a**, CLIP-2dTLIC showing the m⁶A enrichment in HNRNPC-bound RNA regions. Data are mean ± s.d.; n = 3, biological replicates. IP, immunoprecipitation. **b**, HNRNPC-bound RNA regions had higher anti-m⁶A pull-down yield than polyA⁺ RNA. Data are mean ± s.d.; n = 3, biological replicates. **c**, Illustration of the PAR-CLIP-MeRIP protocol. UV, ultraviolet.

d, PAR-CLIP-MeRIP identified an m⁶A residue around the *MALAT1* 2,577 site. **e**, Binding motifs identified with PAR-CLIP-MeRIP peaks. **f**, Density plot showing the positive enrichment at RRACH sites. **g, h**, Validation of two m⁶A-switches by S1/V1 structural probing and filter binding. Data are mean ± s.d.; n = 4, technical replicates. Annotation is the same as in Fig. 1c, d.

residues or nearby regions to modulate protein binding (Fig. 1f). We term this mechanism that regulates RNA–protein interactions through m⁶A-dependent RNA structural remodelling ‘the m⁶A-switch’.

We performed two experiments to determine the global effect of m⁶A-switches on HNRNPC binding. First, *in vivo* crosslinking followed by immunoprecipitation and two-dimensional thin-layer chromatography (CLIP–2dTLC) showed that the m⁶A/A ratio of the HNRNPC-bound RNA regions had a ~6-fold higher m⁶A level than the HNRNPC-bound intact RNA, and a ~3-fold higher m⁶A level than the flow-through RNA (Fig. 2a and Extended Data Fig. 3a). Second, the HNRNPC-bound RNA regions had much higher anti-m⁶A pull-down yield (4.3%) than the polyA⁺ RNA samples (0.5%) using the previously established m⁶A antibody^{13,14} (Fig. 2b). These results indicate a widespread presence of m⁶A residues in the vicinity of HNRNPC-binding sites.

To map the m⁶A sites around HNRNPC-binding sites, we performed PAR-CLIP²⁹ to isolate all HNRNPC-bound RNA regions (input control sample) followed by MeRIP^{13,14} to enrich m⁶A-containing HNRNPC-bound RNA regions (IP sample). Both the input control and IP samples from two biological replicates were sent for RNA sequencing (RNA-seq) (Fig. 2c and Extended Data Fig. 3b, c). This approach, termed PAR-CLIP–MeRIP, identified transcriptome-wide the m⁶A-proximal HNRNPC-binding site, such as the enriched peak around the *MALAT1* 2,577 site (Fig. 2d). Remarkably, HNRNPC PAR-CLIP–MeRIP peaks harboured two consensus motifs, the HNRNPC RBM (U-tracts) and the m⁶A consensus motif GRACH (a subset of RRACH^{13,14}) (Fig. 2e). Both motifs

were located mostly within 50 residues, suggesting transcriptome-wide RRACH–U-tract coupling events within the HNRNPC-binding sites (Extended Data Fig. 4a, b). About 62% of all RRACH–U-tract coupling events within HNRNPC-binding sites are enriched at the RRACH motif (Fig. 2f). Our PAR-CLIP–MeRIP approach identified a total of 39,060 HNRNPC m⁶A-switches that corresponded to m⁶A-modified RRACH–U-tract coupling events at a false discovery rate (FDR) ≤ 5% (Extended Data Fig. 4c). These switches account for ~7% of 592,477 HNRNPC-binding sites identified by PAR-CLIP. The majority (87%) of m⁶A-switches occur within introns (Extended Data Fig. 4d, e), consistent with the literature that HNRNPC is nuclear localized and primarily binds nascent transcripts^{20,23}. We validated two intronic m⁶A-switches in hairpin structures in which m⁶A residues increase the U-tract accessibility and enhance HNRNPC binding by ~3–4 fold (Fig. 2g, h and Extended Data Fig. 5).

To assess the effect of global m⁶A reduction on RNA–HNRNPC interactions, we performed HNRNPC PAR-CLIP experiments in *METTL3* and *METTL14* knockdown cells (Extended Data Fig. 6a). We identified 16,582 coupling events with decreased U-tract–HNRNPC interactions upon *METTL3* and *METTL14* knockdown (*METTL3/L14* knockdown) (Fig. 3a and Extended Data Fig. 6b, c). In total, 2,798 m⁶A-switches identified by PAR-CLIP–MeRIP experiments showed decreased HNRNPC binding upon *METTL3/L14* knockdown (Fig. 3b) and this number is probably an underestimate due to the fact that *METTL3/L14* knockdown reduces the global m⁶A level by only ~30–40% (refs 11, 12). These

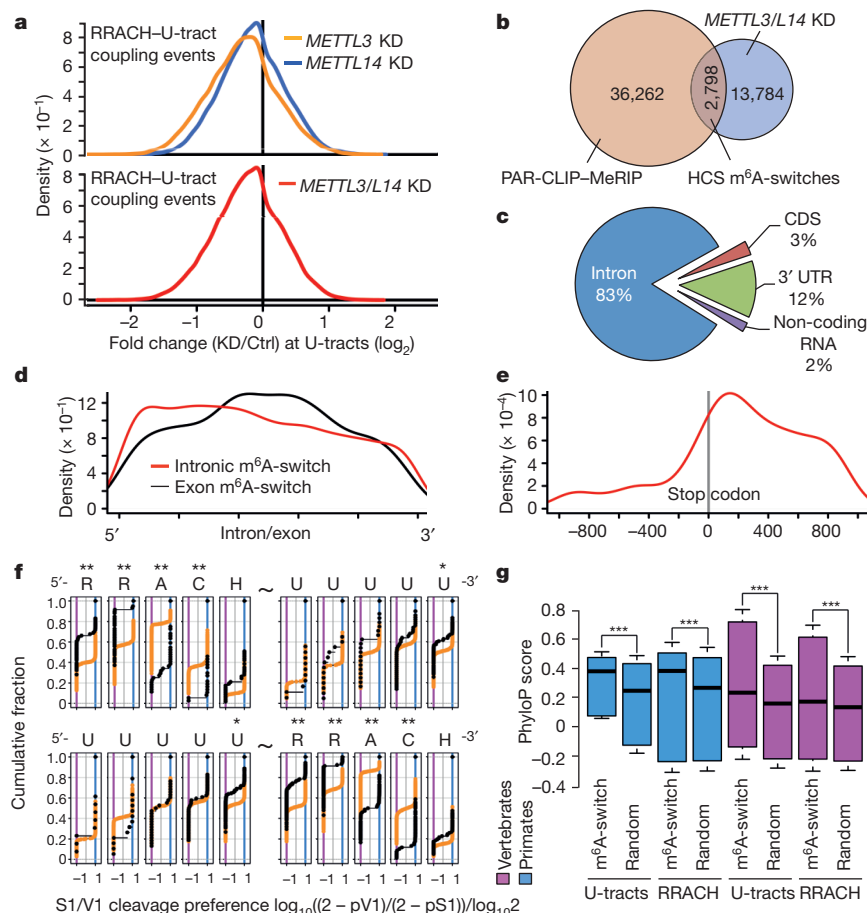


Figure 3 | Global m⁶A reduction decreases HNRNPC binding at m⁶A-switches. **a**, Density plot showing negative enrichment at the U-tracts. KD, knockdown. **b**, Identification of high-confidence (HCS) m⁶A-switches. **c**, Regional distribution of high-confidence m⁶A-switches. CDS, coding sequence. **d**, Density plot showing m⁶A-switch distribution relative to exon/intron boundaries. **e**, m⁶A-switches in coding RNA were enriched in the 3' UTR and near the stop codon. **f**, Cumulative distribution of HCS

m⁶A-switches (black) and control (orange) regarding the S1/V1 cleavage preference (data from ref. 4) at U-tracts and RRACH motif. U-tract can be 3' (top) or 5' (bottom) of the RRACH motif. **P* < 0.05, ***P* < 10^{−4}, Kolmogorov–Smirnov test. **g**, Phylogenetic conservation of high-confidence m⁶A-switches among primates and vertebrates. ****P* < 10^{−16}, Mann–Whitney–Wilcoxon test.

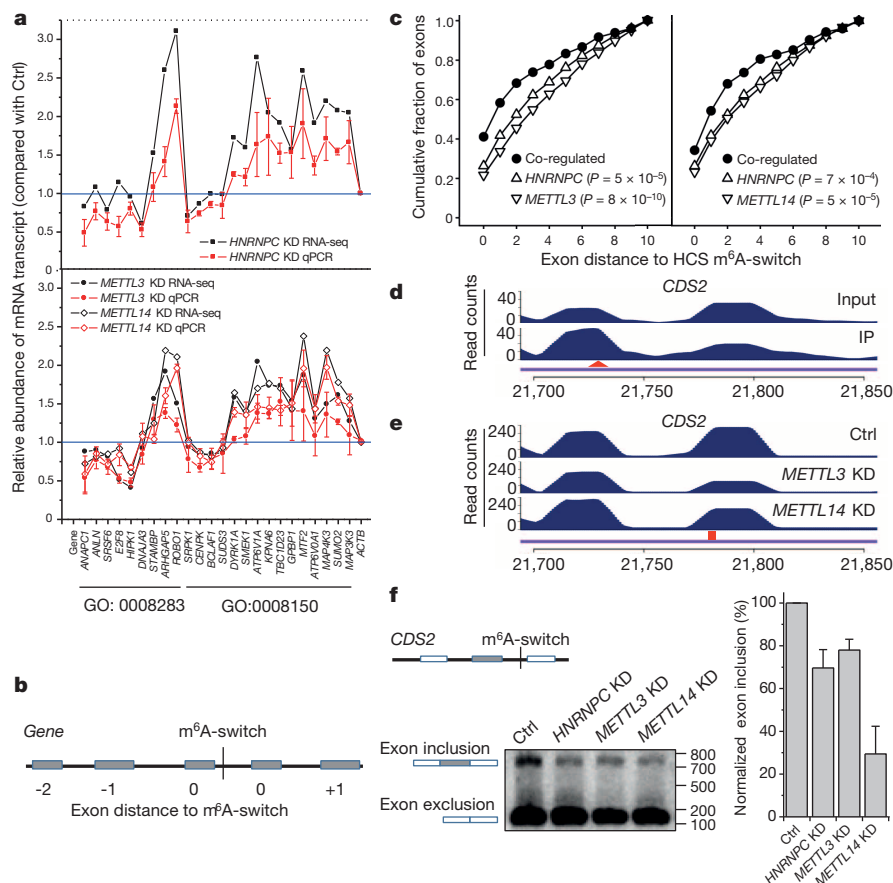
sites composed the high-confidence m⁶A-switches that were used for subsequent analysis.

High-confidence m⁶A-switches are enriched in the introns of coding and non-coding RNAs (Fig. 3c and Extended Data Fig. 6d). Exonic m⁶A-switches are enriched at the middle of exons whereas intronic m⁶A-switches are slightly enriched near the 5' end (Fig. 3d). m⁶A-switches within coding RNAs tend to locate at very long exons (Extended Data Fig. 6e) and are enriched near the stop codon and in the 3' untranslated region (UTR) (Fig. 3e), consistent with the known topology of the human m⁶A methylome in mRNAs^{13,14}. Transcriptome-wide RNA structural mapping^{4–7} on high-confidence m⁶A-switches yielded consistent structural patterns with our three demonstrated m⁶A-switch hairpins (Fig. 3f). The RR residues in the RRACH motif and the 3' U-tract residues show increased structural dynamics in the presence of m⁶A. Besides, m⁶A-switches prefer short RRACH–U-tract inter-motif distances, are not involved in the previously reported inter-U-tract motif patterns and are conserved across species (Fig. 3g and Extended Data Fig. 6f–i).

To reveal the function of m⁶A-switches on RNA biology, we performed polyA⁺ RNA-seq from *HNRNPC*, *METTL3* and *METTL14* knockdown and control cells (Extended Data Fig. 7a). *METTL3/L14* knockdown, which has been shown to decrease HNRNPC binding transcriptome-wide, co-regulated the expression of 5,251 genes with *HNRNPC* knockdown. In comparison, *METTL3/L14* knockdown co-regulated only 24 genes with knockdown of another mRNA-binding protein, HNRNP1 (Extended Data Fig. 7b), which was not enriched in our m⁶A-hairpin

pull down (Fig. 1b). Approximately 45% of 1,815 high-confidence m⁶A-switch-containing genes were co-regulated by *HNRNPC* and *METTL3/L14* knockdown, indicating that m⁶A-switch-regulated HNRNPC binding affects the abundance of target mRNAs. Gene ontology (GO) analysis suggests that m⁶A-switch-regulated gene expression may influence 'cell proliferation' and other biological processes (Extended Data Fig. 7c). The m⁶A-switch-regulated expression of genes within these GO categories was validated by quantitative polymerase chain reaction (qPCR) (Fig. 4a and Extended Data Fig. 7d–g). We also found that *HNRNPC*, *METTL3* and *METTL14* knockdown decreased the cell proliferation rate to similar extents (Extended Data Fig. 7h).

Besides the mRNA abundance level changes, we also observed splicing pattern changes within high-confidence m⁶A-switch-containing transcripts by testing the differential exon usage in RNA-seq data (DEXSeq)³⁰. *HNRNPC* knockdown co-up/downregulated 131/127 exons with *METTL3* knockdown and 130/115 exons with *METTL14* knockdown. These co-regulated exons occur more frequently in the vicinity of m⁶A-switches than non-co-regulated exons (Fig. 4b, c), indicating that m⁶A-switches tend to regulate splicing events at nearby exons. We investigated the splicing pattern at two exons with neighbouring m⁶A-switches: the PAR-CLIP–MeRIP and *METTL3/L14* knockdown data confirmed the HNRNPC-binding signature at the m⁶A-switch site neighbouring these exons; and *HNRNPC* and *METTL3/L14* knockdown co-inhibited exon inclusion in both cases (Fig. 4d–f and Extended Data Fig. 8b–f). Besides, we identified 155 genes with multiple m⁶A-switches exhibiting more



than two splice variants, and 221 m⁶A-switch-containing genes with differentially expressed splice variants in *HNRNPC* and *METTL3/L14* knockdown samples. Further analysis suggested that m⁶A-switches have an effect on intron exclusion (Extended Data Fig. 8g). Consistent with previous reports about splicing regulation by both *HNRNPC* and m⁶A^{13,19,20,23}, our results indicate that m⁶A functions as an RNA structure remodeler to affect mRNA maturation through interference with post-transcriptional regulator binding activities.

We demonstrated that post-transcriptional m⁶A modifications could modulate the structure of coding and non-coding RNAs to regulate RNA–HNRNPC interactions, thus influencing gene expression and maturation in the nucleus. It is possible that m⁶A could also recruit additional accessory factors, such as the YTH domain proteins, which can directly recognize m⁶A—as previously reported¹⁵—to destabilize the RNA structure and facilitate HNRNPC binding. Besides HNRNPC, m⁶A-switches may regulate the function of many other RNA-binding proteins through modulating the RNA-structure-dependent accessibility of their RBMs. Our work indicates widespread m⁶A-induced mRNA and lncRNA structural remodelling that affects RNA–protein interactions for biological regulation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 November; accepted 14 January 2015.

- Antson, A. A. Single-stranded-RNA binding proteins. *Curr. Opin. Struct. Biol.* **10**, 87–94 (2000).
- Dreyfuss, G., Kim, V. N. & Kataoka, N. Messenger-RNA-binding proteins and the messages they carry. *Nature Rev. Mol. Cell Biol.* **3**, 195–205 (2002).
- Ray, D. *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177 (2013).
- Wan, Y. *et al.* Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**, 706–709 (2014).
- Ding, Y. *et al.* *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**, 696–700 (2014).
- Kertesz, M. *et al.* Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467**, 103–107 (2010).
- Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. & Weissman, J. S. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures *in vivo*. *Nature* **505**, 701–705 (2014).
- Bokar, J. A. in *Fine-Tuning of RNA Functions by Modification and Editing* (ed. Grosjean, H.) 141–178 (Springer, 2005).
- Jia, G. *et al.* N⁶-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chem. Biol.* **7**, 885–887 (2011).
- Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18–29 (2013).
- Liu, J. *et al.* A METTL3–METTL14 complex mediates mammalian nuclear RNA N⁶-adenosine methylation. *Nature Chem. Biol.* **10**, 93–95 (2014).
- Wang, Y. *et al.* N⁶-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nature Cell Biol.* **16**, 191–198 (2014).
- Dominissini, D. *et al.* Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* **485**, 201–206 (2012).
- Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
- Wang, X. *et al.* N⁶-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).
- Fustin, J. M. *et al.* RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* **155**, 793–806 (2013).
- Schwartz, S. *et al.* High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* **155**, 1409–1421 (2013).
- Batista, P. J. *et al.* m⁶A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707–719 (2014).
- Zhao, X. *et al.* FTO-dependent demethylation of N⁶-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* **24**, 1403–1419 (2014).
- König, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Struct. Mol. Biol.* **17**, 909–915 (2010).
- McCloskey, A., Taniguchi, I., Shinmyozu, K. & Ohno, M. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* **335**, 1643–1646 (2012).
- Rajagopalan, L. E., Westmark, C. J., Jarzembowski, J. A. & Malter, J. S. hnRNP C increases amyloid precursor protein (APP) production by stabilizing APP mRNA. *Nucleic Acids Res.* **26**, 3418–3423 (1998).
- Zarnack, K. *et al.* Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* **152**, 453–466 (2013).
- Cieniková, Z. *et al.* Structural and mechanistic insights into poly(uridine) tract recognition by the hnRNP C RNA recognition motif. *J. Am. Chem. Soc.* **136**, 14536–14544 (2014).
- Liu, N. *et al.* Probing N⁶-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* **19**, 1848–1856 (2013).
- Krecic, A. M. & Swanson, M. S. hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* **11**, 363–371 (1999).
- Görlach, M., Burd, C. G. & Dreyfuss, G. The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J. Biol. Chem.* **269**, 23074–23078 (1994).
- Kierzek, E. & Kierzek, R. The thermodynamic stability of RNA duplexes and hairpins containing N⁶-alkyladenosines and 2-methylthio-N⁶-alkyladenosines. *Nucleic Acids Res.* **31**, 4472–4480 (2003).
- Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
- Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).

Acknowledgements This work is supported by National Institutes of Health EUREKA GM088599 (C.H. and T.P.), and by K01HG006699 (Q.D.). We thank all members of the Pan and He laboratories for comments and discussions. We also thank Y. C. Leung, G. Perdrietz, Y. Pigli, J. Yue, J. Liu, Y. Yue, K. Chen and M. Yu for technical assistance. C.H. is an Investigator of the Howard Hughes Medical Institute. M.P. was a Natural Sciences and Engineering Research Council of Canada postdoctoral fellow.

Author Contributions N.L., G.Z. and M.P. designed and performed experiments, and analysed data. Q.D. synthesized all RNA oligonucleotides. N.L., M.P. and T.P. conceived the project. N.L. and T.P. wrote the paper with input from C.H. and M.P.

Author Information RNA-seq data have been deposited in the Gene Expression Omnibus under accession number GSE56010. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.P. (marc.parisien@mcgill.ca) or T.P. (taopan@uchicago.edu).

METHODS

Mammalian cell culture, siRNA knockdown and western blot. Human cervical cancer cell line HeLa (CCL-2) and embryonic kidney cell line HEK293T (CRL-11268) were obtained from the American Type Culture Collection (ATCC) and were cultured under standard conditions. Control short interfering RNA (siRNA) (1027281, Qiagen), *METTL3* siRNA (SI04317096, Qiagen), *METTL14* siRNA (SI04317096, Qiagen) or *HNRNPC* siRNA (10620318, Invitrogen) were transfected into HEK293T cells at a concentration of 40 nM using lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. Cells were collected 48 h after the transfection, shock-frozen in liquid nitrogen, and stored at -80°C for further studies. Western blot analysis using *METTL3*- (HPA038002, Sigma), *METTL14*- (HPA038002, Sigma), *HNRNPC*- (sc-32308, Santa Cruz) and *GAPDH*- (A00192-40, Genescript) specific antibodies was performed under standard procedures. Blotting membranes were stained by ECL-prime (RPN2232, GE Healthcare) and visualized by a digital imaging system (G: BOX, SYNGENE). All synthetic oligonucleotides were synthesized by Q.D.

Gel shift, RNA pull-down and filter-binding assays. HeLa nuclear extracts were isolated using the NE-PER Nuclear and Cytoplasmic Extraction Reagents (78833, Thermo Scientific) according to the manufacturer's instructions. The purified radioactively labelled RNA oligonucleotides were refolded by heating at 90°C for 1 min, then at 30°C for 5 min. Three microlitres HeLa nuclear extract and 6 μl refolded RNA were incubated at room temperature for 30 min and then at 4°C for 2 h. Each sample was mixed with 1 μl 50% glycerol, separated on an 8% native $1\times$ TBE gel, and visualized by phosphorimaging using the Personal Molecular Imager (Bio-Rad).

The *in vitro* pull-down assay was performed as described¹³. The eluted protein samples were separated on 4–12% polyacrylamide Bis-Tris gels (NP0321BOX, Invitrogen) and stained with SYPRO-Ruby (S12000, Invitrogen) according to the manufacturer's instructions. Protein in gel slices or the entire pulled-down protein samples were digested with trypsin and identified using Liquid chromatography-tandem mass spectrometry by the Donald Danforth Plant Science Center (Washington University). The RNA oligonucleotides used in Fig. 1f were: 2,577–U, 5'-A ACUUA AUGUUUUGCAUUGGUCUUUGAGUUA-Biotin; 2,577–CC-A, 5'-A ACUUA AUGUCCUUGCAUUGGACUUUGAGUUA-Biotin; 2,577–CC-m⁶A, 5'-AACUUA AUGUCCUUGCAUUGGm⁶ACUUUGAGUUA-Biotin.

The full-length HNRNPC1 protein was purified and the *in vitro* ultraviolet cross-linking assay was performed as previously described²³. Filter-binding assays were performed as previously described²⁴.

CLIP-2dTLC. HEK293T cells at 70–80% confluency were ultraviolet irradiated with 400 mJ cm^{-2} at 254 nm, and harvested by centrifuging at 4,000 r.p.m. for 3 min at 4°C (with centrifugation rotor 75003524, Fisher Scientific). The pellet of cross-linked cells were resuspended in 1 ml lysis buffer ($1\times$ PBS, 0.1% SDS, 1% Nonidet P-40, 0.5% sodium deoxycholate, protease inhibitor cocktail and RNase inhibitor) and incubated on ice for 4 h. Cell lysate was isolated by centrifuging at 3,000 r.p.m. for 5 min and pre-blocked with 50 μl protein A beads in 300 μl lysis buffer. Another 50 μl protein A beads (Invitrogen) were incubated with 8 μg corresponding antibodies for 4 h at room temperature, and then mixed with the pre-blocked cell lysate at 4°C overnight. The beads were washed three times with 1 ml wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl₂, 0.2% Tween-20), three times with 1 ml high-salt buffer ($5\times$ PBS, 0.1% SDS, 1% Nonidet P-40, 0.5% sodium deoxycholate), and three times with 1 ml wash buffer. The beads were resuspended in 1 ml wash buffer, and divided into $2\times$ 500 μl in two separate tubes. One tube was incubated with 200 μl RNase T1/A mixture at room temperature for 1 h. The other tube was incubated with 200 μl nuclease-free water at room temperature for 1 h. The beads were washed three times with 1 ml high-salt buffer, and three times with 1 ml wash buffer. Cross-linked RNA was eluted from beads by incubating with 200 μl RNA elution buffer (100 mM Tris-HCl pH 7.4, 10 mM EDTA, 1% SDS) containing 2 mg ml⁻¹ proteinase K at 50°C for 30 min followed by phenol/chloroform extraction. The RNA pellet was dissolved in 7 μl nuclease-free water containing 1 μl RNase T1 (200 U), heated at 65°C for 2 min, and incubated at 37°C for 30 min. The T1-digested RNA fragments were labelled upon adding 2 μl T4 PNK mix (4.5 U μl^{-1} T4 PNK, 600 Ci mmol⁻¹ [γ -³²P] ATP, $5\times$ PNK buffer) and incubation at 37°C for 30 min. Unreacted [γ -³²P] ATP was removed using Illustra MicroSpin G-25 columns. The eluted RNA was digested with 1 μl (1 U μl^{-1}) nuclease P1 at 37°C for 1 h. Samples were spotted on cellulose TLC plate and 2dTLC was run as described²⁵ using isobutyric acid: 0.5 M NH₄OH (5:3, v/v) as the first dimension and isopropanol:HCl:water (70:15:15, v/v/v) as the second dimension.

RNA structural probing and RNA terminal truncation. The synthetic RNA oligonucleotides were 5'-end-labelled with γ -³²P-ATP by T4 PNK (70031, Affymetrix), gel purified, and re-folded. Structural probing assay with RNase T1, nuclease S1 and RNase V1 was performed as previously described²⁵. Note that 3'-end-labelled *HNRNP1* oligonucleotides were used for the RNA structural probing assay shown in Fig. 2g.

CMCT RNA structural probing assay was performed as reported³¹. RNA refolding: 3 pmol RNA was annealed in 50 mM potassium borate (pH 8) by heating at 90°C for 1.5 min then incubation at room temperature for 3 min.

RNA terminal truncation assay was carried out as previously reported³². RNA samples were first alkaline-hydrolysed as in the RNA structural probing assay, and then incubated with HNRNPC1 protein in the same conditions as in the filter binding assay. The RNA-protein complexes were then loaded onto filter papers and washed twice with chilled binding buffer. Air-dry filters and RNA samples were then extracted from the filters and loaded onto denaturing gel as in the RNA structural probing assay.

PAR-CLIP and PAR-CLIP-MeRIP. PAR-CLIP procedures were performed as previously reported²⁹ with the following modification. HEK293T cells in 15-cm plates treated following normal PAR-CLIP procedures were lysed and digested with a combination of RNase I (Ambion, AM2295, 15 μl 1/50 diluted with H₂O) and Turbo DNases (2 μl) for 3 min at 37°C , shaking at 1,100 r.p.m. The lysate was then immediately cleared by spinning at 14,000 r.p.m., 4°C for 30 min, and placed on ice for further use. HNRNPC-binding sites were identified by PARalyzer v.1.1 (ref. 33) with default settings.

PAR-CLIP-MeRIP experiment applied m⁶A-antibody immunoprecipitation^{13,14,34} to the HNRNPC PAR-CLIP RNA samples. The HNRNPC PAR-CLIP RNA sample was incubated with m⁶A-specific antibody (202003, SYSY), RNase inhibitor (80 units, Sigma-Aldrich), human placental RNase inhibitor (NEB) in 200 μl $1\times$ IP buffer (50 mM Tris-HCl pH 7.4, 750 mM NaCl and 0.5% (v/v) Igepal CA-630) at 4°C for 2 h under gentle shaking conditions. For each PAR-CLIP-MeRIP experiment, 20 μl protein A beads (Invitrogen) were washed twice with 1 ml $1\times$ IP buffer, blocked with 2 h incubation with 100 μl $1\times$ IP buffer supplemented with bovine serum albumin (BSA) (0.5 mg ml⁻¹), RNasin and human placental RNase inhibitor, and then washed twice with 100 μl $1\times$ IP buffer. The pre-blocked protein A beads were then combined with the prepared immuno-reaction mixture and incubated at 4°C for 2 h, followed by three washes with 100 μl $1\times$ IP buffer. After that, the RNA was eluted by 1 h incubation with 20 μl elution buffer ($1\times$ IP buffer and 6.7 mM m⁶A, Sigma-Aldrich) under gentle shaking conditions, and purified by ethanol precipitation. The purified RNA sample (IP) as well as the input PAR-CLIP RNA sample (input control) were used for library construction by Truseq small RNA sample preparation kit (Illumina).

Libraries were prepared using TruSeq Small RNA Sample Preparation Kit (RS-200-0012, Illumina) according to the manufacturer's instructions, and then sequenced by Illumina HiSeq2000 with single-end 50-bp read length. The control and IP samples from PAR-CLIP-MeRIP experiments (same case for the control and knockdown samples from *METTL* knockdown experiments) were sequenced together in one flowcell on two lanes, and the reads from two lanes of each sample were combined for remaining analysis. The raw sequencing data were trimmed using the Trimmomatic computer program v.0.30 (ref. 35) to remove adaptor sequences, and mapped to the human genome version hg19 by Bowtie 1.0.0 (ref. 36) without any gaps and allowed for at most two mismatches.

Detection of PAR-CLIP-MeRIP peaks and differential PAR-CLIP peaks. The raw read counts of the biological replicates confirmed the reproducibility between replicates (Extended Data Fig. 9), and replicates were combined for subsequent analysis. For each genomic site, we calculated the average read counts within an 11-nucleotide window centred at that site, as the normalized read counts for that site. This normalization smoothed the raw mapping curves, and facilitated identification of peaks within each mapping cluster. To correct for changes in sequencing depth or expression levels between samples, we then normalized the read counts at each genomic site to the total number of read counts on the respective gene. The above defined double-normalization procedures enabled precise identification of changes in the mapping reads at specific genomic locations by directly comparing the normalized read counts between samples. No read counts in the intergenic region were compared between samples, because the transcription boundaries are not defined at this region and the intergenic read counts cannot be normalized to correct changes for transcript expression.

Detection of PAR-CLIP-MeRIP peaks involves comparing the read counts of the IP sample with that of the control (Ctrl) sample as follows: (1) we identified all peaks within HNRNPC-binding sites in the IP sample; (2) we performed transcriptome-wide scanning to compare read counts of each identified peak in (1) with read counts at the same genomic locations in the Ctrl sample to calculate the fold change score, $\text{score} = \log_2(H_{\text{IP}}/H_{\text{Ctrl}})$. The score threshold was set to be 1, corresponding to a twofold increase compared with control.

The detection of decreased HNRNPC-binding sites involved comparing HNRNPC occupancies in the *METTL* knockdown (KD) sample with that in the control as follows: (1) we identified all peaks within HNRNPC-binding sites in the *METTL* knockdown sample; (2) we performed transcriptome-wide scanning to compare read counts of each identified peak in (1) with read counts at the same genomic locations in control to calculate the fold change score, $\text{score} = \log_2(H_{\text{KD}}/H_{\text{Ctrl}})$.

The score threshold was set to be -1 , corresponding to a twofold decrease compared with control.

Identification of enriched motifs and HNRNPC m⁶A-switches. To identify enriched motifs, we first sorted the 12,998 HNRNPC PAR-CLIP-MeRIP peaks (with IP/input enrichment ≥ 2) by the T-to-C mutation frequency. We then chose the top 4,500 peaks with the highest T-to-C mutation frequency for motif analysis using FIRE³⁷ with default RNA analysis parameters. The top two enriched motifs are the GRACH and the U-tract motif. We also used the top 1,024 and 2,048 peaks for motif analysis, yielding the same motif results as the top 4,500 peaks.

To identify transcriptome-wide HNRNPC m⁶A-switches, we first searched for all coupling events within 50 nucleotides between U₅ and RRACH motifs, with the U₅ motif located within HNRNPC-binding sites. For PAR-CLIP-MeRIP samples, the fold-change score E at the RRACH motif was calculated for each coupling event. Also, the P value for each coupling event was calculated as described³⁸. Then, we generated the π value, $\pi = E \cdot (-\log_{10} P)$, as one comprehensive parameter to pick meaningful genomic loci³⁹. HNRNPC m⁶A-switches identified from PAR-CLIP-MeRIP experiments should fulfill the following requirements: (1) read counts at both the control and IP sample ≥ 5 ; (2) π value ≥ 0.627 , corresponding to FDR $\leq 5\%$.

For *METTL* knockdown samples, the fold-change score at the U-tract motifs was calculated for each coupling event. HNRNPC m⁶A-switches identified from *METTL3/L14* knockdown samples should fulfill the following requirements: (1) read counts at both the control and knockdown sample ≥ 5 ; (2) π value ≤ 0.627 , corresponding to FDR $\leq 5\%$.

Distribution of HNRNPC m⁶A-switches. Pie charts illustrating the distribution within each segment were made using the following hierarchy: intron > ncRNA > 3' UTR > 5' UTR > CDS > intergenic. To plot the distribution of HNRNPC m⁶A-switches in their respective localized segments (such as intron, exon, 3' UTR, CDS, 5' UTR), we first identified the distance between each m⁶A-switch and the 5' end of the respective segment. This distance was then divided by the length of that segment to determine a percentile where this m⁶A-switch fell, and then this specific percentile bin was incremented. Following this approach, we obtained the distribution pattern of all m⁶A-switches within each segment.

RNA-seq. RNA-seq experiments were performed on two replicate RNA samples from *HNRNPC*, *METTL3*, *METTL14* knockdown as well as control HEK293T cells (48 h after transfection). Total RNA samples were extracted according to the RNeasy Plus Kit (catalogue no. 74104, Qiagen). Libraries were prepared according to the TruSeq Stranded mRNA LT Sample Prep Kit (catalogue no. RS-122-9005DOC). Knockdown and control samples were sequenced together in one flowcell on four lanes, respectively. All samples were sequenced by Illumina HiSeq 2000 with pair end 100-bp read length. The reads from the four lanes of each sample were combined for all analyses. The RNA-seq data were mapped using the splice-aware alignment algorithm TopHat v.1.1.4 (ref. 40) based on the following parameters: tophat -num-threads 8 -mate-inner-dist 200 -solexa-quals -min-isoform-fraction 0 -coverage-search-segment-mismatches 1. Gene expression level changes were analysed using cuffdiff⁴¹. Differential splicing was determined using DEXSeq⁴⁰ based on Cufflinks-predicted, non-overlapping exons. To compare with a different mRNA-binding protein, the RNA-seq data from *HNRNP* knockdown HEK293T cells (GEO34995 data set⁴²) were analysed.

GO, evolutionary conservation, graphic and statistical analyses. GO enrichment analysis was applied on the co-regulated high-confidence m⁶A-switch-containing genes, against all high-confidence m⁶A-switch-containing genes as background, using GOrilla⁴³.

Phylogenetic conservation analysis was performed by comparing PhyloP scores at the U-tract motif and RRACH motif for HNRNPC m⁶A-switches to those of randomly selected sequences. The PhyloP scores were accessed from the precompiled PhyloP scores⁴⁴ (<ftp://hgdownload.soe.ucsc.edu/goldenPath/hg19/phyloP46way/>) under both primate and vertebrate categories. P values were evaluated using the Mann-Whitney-Wilcoxon test, $***P < 10^{-16}$. For the U-tract motifs, we collected all U-tracts (5 × U₅) across all chromosomes and randomly selected 10,000 sites among the 38,561,577 sites of our census. The random selection was done separately for primates and for vertebrates. For the RRACH motif, we also collected all RRACH sites across all chromosomes and randomly selected 10,000 sites among the 78,815,225 sites of our census. Here, too, the random selection was done separately for primates and vertebrates.

Sequence logos were generated using the WebLogo package. The R statistical package was used for all statistical analyses (unless stated otherwise).

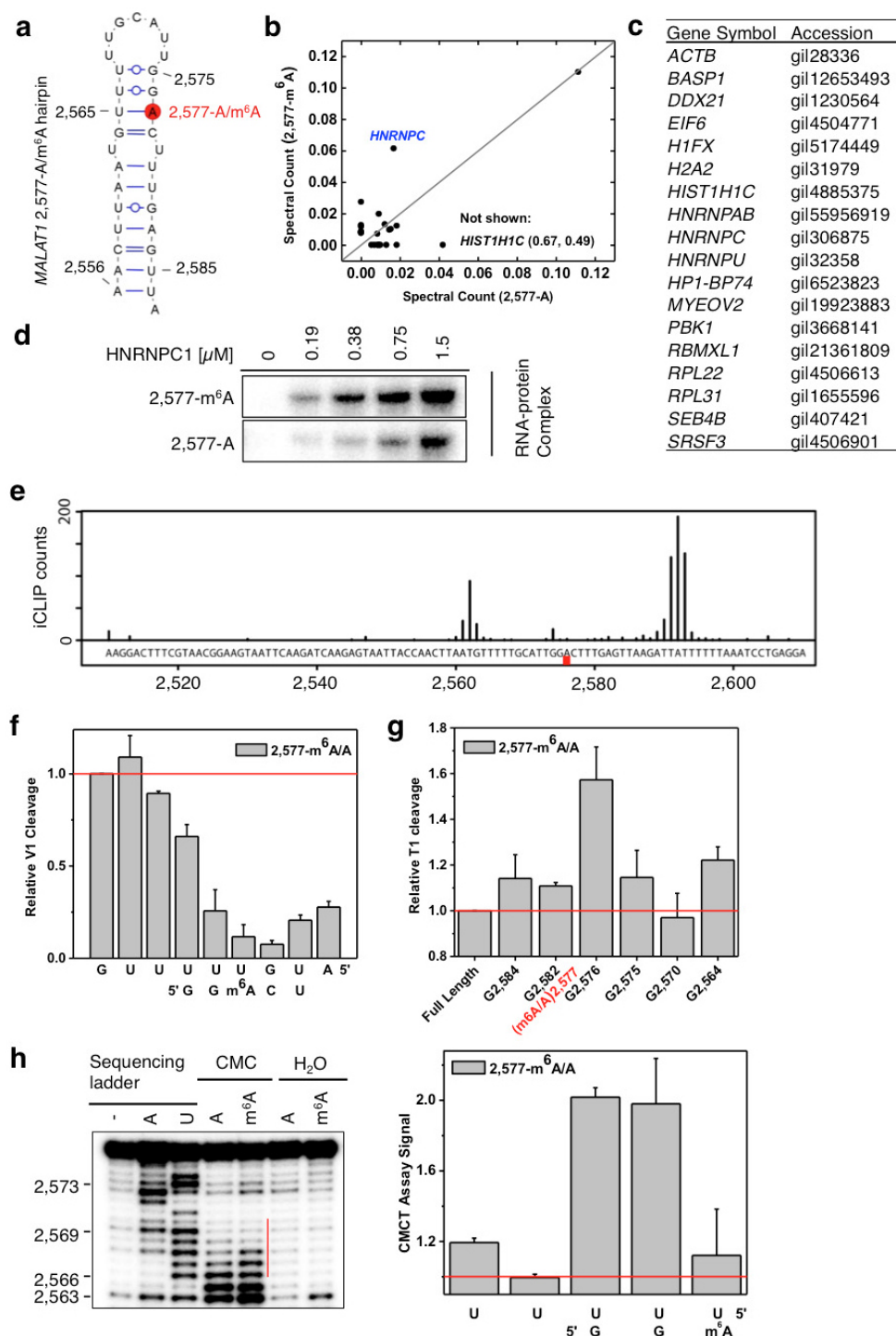
Cell proliferation analysis. HEK293T cells were transfected with si-control, si-*HNRNPC*, si-*METTL3* and si-*METTL14* RNAs. After transfection, the numbers of cells were counted at 0, 24, 48 and 72 h, as described previously⁴⁵. Three independent experiments were performed and growth curves were plotted to test the effects on cell proliferation.

RT-PCR quantitation. Total RNA samples were extracted from HEK293T cells and reverse transcribed using SuperScript III First-Strand Synthesis System (Life

Technologies, catalogue no. 18080-051). In order to validate the splicing changes identified from our RNA-seq data, we performed RT-PCR measurements using Thermo Scientific Taq DNA Polymerase under the following conditions: 95 °C for 3 min, 30 cycles of 95 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min, and then finally 72 °C for 10 min. For the target alternate exon, we designed and used primers annealing to both neighbouring constitutive exons. The PCR products were separated on 1.2% agarose gel and ethidium bromide stained. In order to validate the gene expression level changes identified from our RNA-seq data, we performed qRT-PCR measurements using Power SYBR Green PCR Master Mix (Life Technology, catalogue no. 4367659) under the following conditions: 50 °C for 3 min followed by 95 °C for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 1 min, and then 40 °C for 1 min and 95 °C for 15 s and finally 60 °C for 30 s.

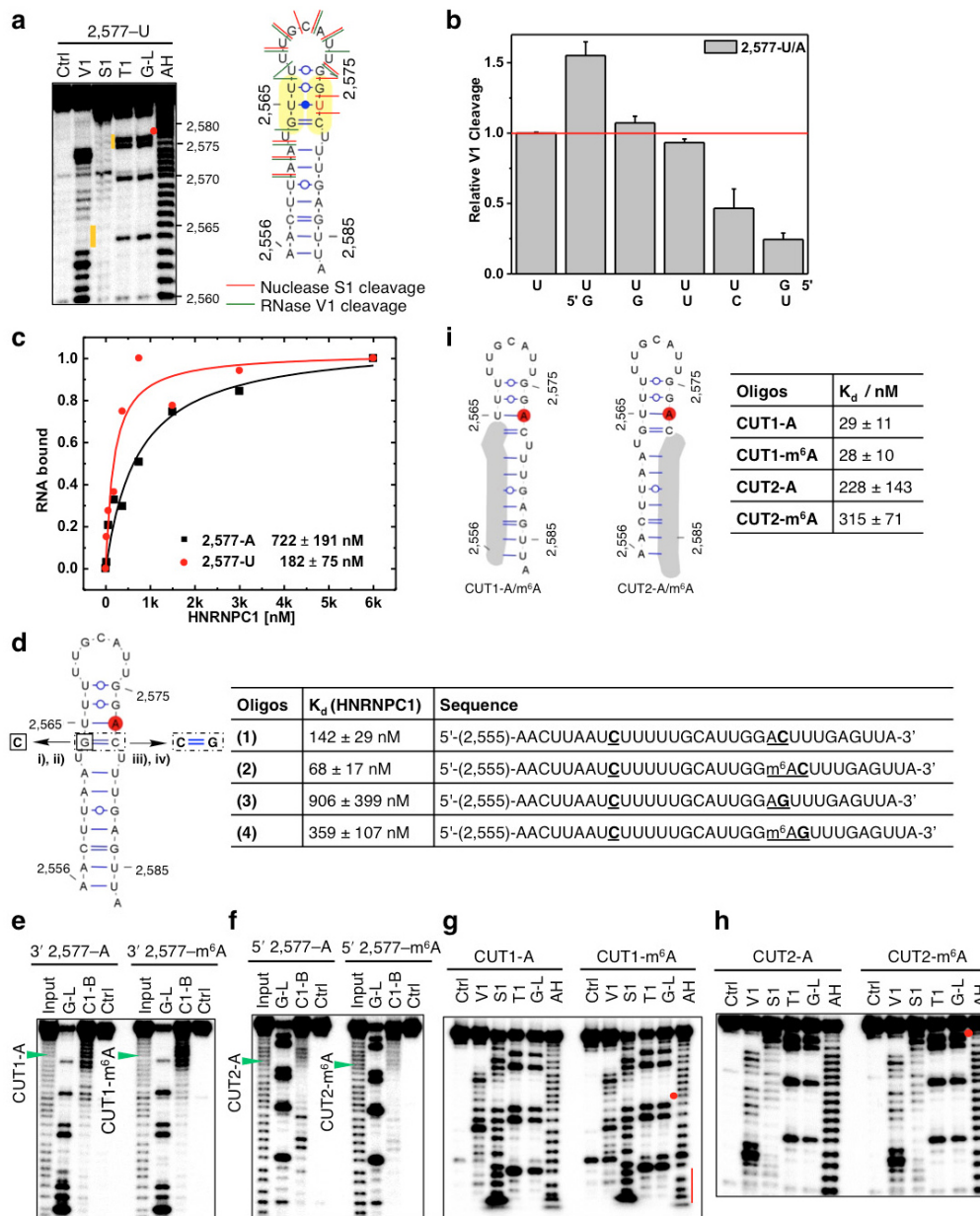
The primer sequences are as follows (listed as gene name: forward primer; reverse primer). *ANAPC1*: TGCCAAAAGAAATAGCAGTTCAG; TGCCAAAAGAAATAGCAGTTCAG; *ANLN*: GCCAGGCGAGAGAATCTTCA; GGCTGCTGGTTACTTGCTTC; *SRSF6*: ACAAGGAACGAACAAATGAGGG; GCTCCAGAGTAAGATCGCCTAT; *E2F8*: ACCCAAGCTCAGCCATTGTA; GAGTCATAGTTGGTGGCCCT; *HIPK1*: CCAGTCAGCTTTGTACCCATC; TTGAAACGCAGGTGGACATA; *DNAJA3*: CCCTTTCATTTGTACTGCCTCC; TGATCTCTTTCTGGCTGGCA; *STAMPB*: GTTCTCATCCCCAAGCAAAG; ATCCAGCCCAGTGTGATGA; *ARHGAP5*: GCGGATTCCATTGACCTCC; GTGCCCCGTGGTGAATGAAT; *ROBO1*: TTTGGGCTTCTGCGTATGTT; GAGGGGTACTGGA GACAGCA; *SRPK1*: CCCTGAGAAGAGAGCCACTG; ACCCTGAAAAGGGAAGAGGA; *CENPK*: AAGGCTAAAAATTCACAAAGCA; TCCATATCTTTCCACATTTCTTCA; *BCLAF1*: TCCTGAAAGGTCTGGGTCTG; TCCTGAAAGGTCTGGGTCTG; *SUDS3*: TGCCTGGGGTCTGTATTTTC; CAGTTCAAGCGA GGGGAAGTC; *DYRK1A*: CTTCAGCATGCAAACTTCA; GGCAGAAACCTGTGTGTCAC; *SMK1*: TTGAAGGACTGCACCACTG; CCTGTGTTTCTGTTGTG; *ATP6V1A*: AAGCATTTCCCTCTGTCAA; CTGCCAGGTCTTCTTCTTCC; *KPNA6*: CCCTGTGTTGATCGAAATCC; GATCTGCTCAGGGGTTCTCTC; *TBC1D23*: GGTGAATCTCCTAATGGCTCA; CGATCCACAGGAGTTGATGT; *GPBP1*: CGTCATTGAATTTTGAAGAAGCA; TTAGGACGCCCCATAGCAGA; *MTF2*: GTCTGATTTGGTTCTGGT; CTGCAGGAAAGGCAACCTTA; *ATP6V0A1*: TCCGTGTCTGGTTCATCAAA; TCTGAGTGCAAACTGGATGG; *MAP4K3*: TCTTCATACCACAGGAAATGC; AACAGGTTTGTGTGGGGGTA; *SUMO2*: TTCTTTCATTTCCCCCTTCC; TATTTTCCCCATCCGCTCT; *MAP3K3*: CAGTTCCTCTCCCCACTCTG; GACAGAGAGGTGCTGCTC; *CDS2*: CGATTTTCCAGGATGACAG; GAAAGGGCCCTATTGAGGAC; *YTHDF2*: ACTTGAGTCCACAGGCAAGG; AAGCAGCTTACCCAAAGAA.

- Ehresmann, C. et al. Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**, 9109–9128 (1987).
- Peterson, E. T., Pan, T., Coleman, J. & Uhlenbeck, O. C. *In vitro* selection of small RNAs that bind to *Escherichia coli* phenylalanyl-tRNA synthetase. *J. Mol. Biol.* **242**, 186–192 (1994).
- Corcoran, D. L. et al. PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. *Genome Biol.* **12**, R79 (2011).
- Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N. & Rechavi, G. Transcriptome-wide mapping of N⁶-methyladenosine by m⁶A-seq based on immunocapturing and massively parallel sequencing. *Nature Protocols* **8**, 176–189 (2013).
- Lohse, M. et al. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* **40**, W622–W627 (2012).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Elemento, O., Slonim, N. & Tavazoie, S. A universal framework for regulatory element discovery across all genomes and data types. *Mol. Cell* **28**, 337–350 (2007).
- Ouyang, Z., Snyder, M. P. & Chang, H. Y. SeqFold: genome-scale reconstruction of RNA secondary structure integrating high-throughput sequencing data. *Genome Res.* **23**, 377–387 (2013).
- Xiao, Y. et al. A novel significance score for gene selection and ranking. *Bioinformatics* **30**, 801–807 (2014).
- Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
- Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562–578 (2012).
- Huelga, S. C. et al. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep.* **1**, 167–178 (2012).
- Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
- Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).
- Yang, F., Yi, F., Han, X., Du, Q. & Liang, Z. MALAT-1 interacts with hnRNP C in cell cycle regulation. *FEBS Lett.* **587**, 3175–3181 (2013).



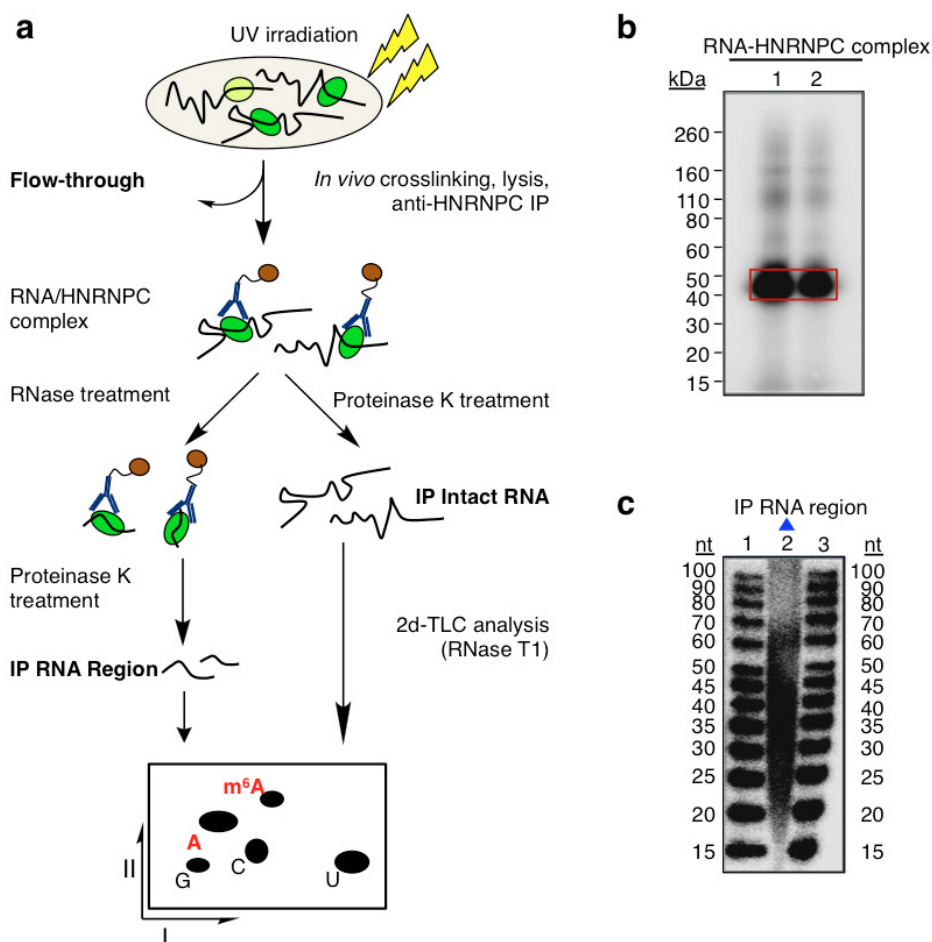
Extended Data Figure 1 | m⁶A increases the accessibility of the U-tract to enhance HNRNPC binding. **a**, Secondary structure of the *MALAT1* hairpin with m⁶A methylation at the 2,577 site shown in red²⁵. Nucleotide position numbers correspond to their locations along the human *MALAT1* transcript (NCBI accession NR_002819). **b**, RNA pull down showing that HNRNPC preferably binds methylated RNA. **c**, The list of proteins with identified peptides by mass spectrometry in **b**. **d**, Recombinant HNRNPC1 binds more strongly with that *MALAT1* 2,577-m⁶A hairpin compared with the unmethylated hairpin, as determined by an *in vitro* ultraviolet crosslinking assay²³. **e**, HNRNPC shows binding around the 2,577-A site along *MALAT1* *in vivo*, as determined by previously published HNRNPC iCLIP data²⁰. The underlying genomic sequence is shown at the bottom with a red square marking the 2,577-m⁶A site. The slight shift of the iCLIP signal to upstream of the U-tract-binding site is probably due to the steric hindrance of the peptide fragment remaining on RNA, which can cause reverse transcription to

terminate more than one nucleotide upstream of the crosslink site²⁰. **f**, Quantification of the RNase V1 cleavage signal for the U-tract region from the RNA structural mapping assay in Fig. 1e. To correct for sample loading difference, each band signal was normalized to the band signal of the immediate 3' residue to the U-tract. Data are mean \pm s.d.; $n = 3$, technical replicates. **g**, Quantitative analysis of the RNase T1 cleavage signal from the RNA structural mapping assay in Fig. 1e. An increased RNase T1 cleavage signal (single-strand specific and cleavage after guanosines) was observed due to the surrounding m⁶A residue. To correct for sample loading difference, the ratio for each band signal among all bands in each lane was calculated. Relative T1 cleavage = (m⁶A_{native}/m⁶A_{denature})/(A_{native}/A_{denature}). $n = 2$, technical replicates. **h**, Quantitative CMCT mapping showing increased signals for the U-tract bases around the U base-pairing with m⁶A. Quantitation of band signals within the U-tract region is shown on the right. Data are mean \pm s.d.; $n = 4$, technical replicates.



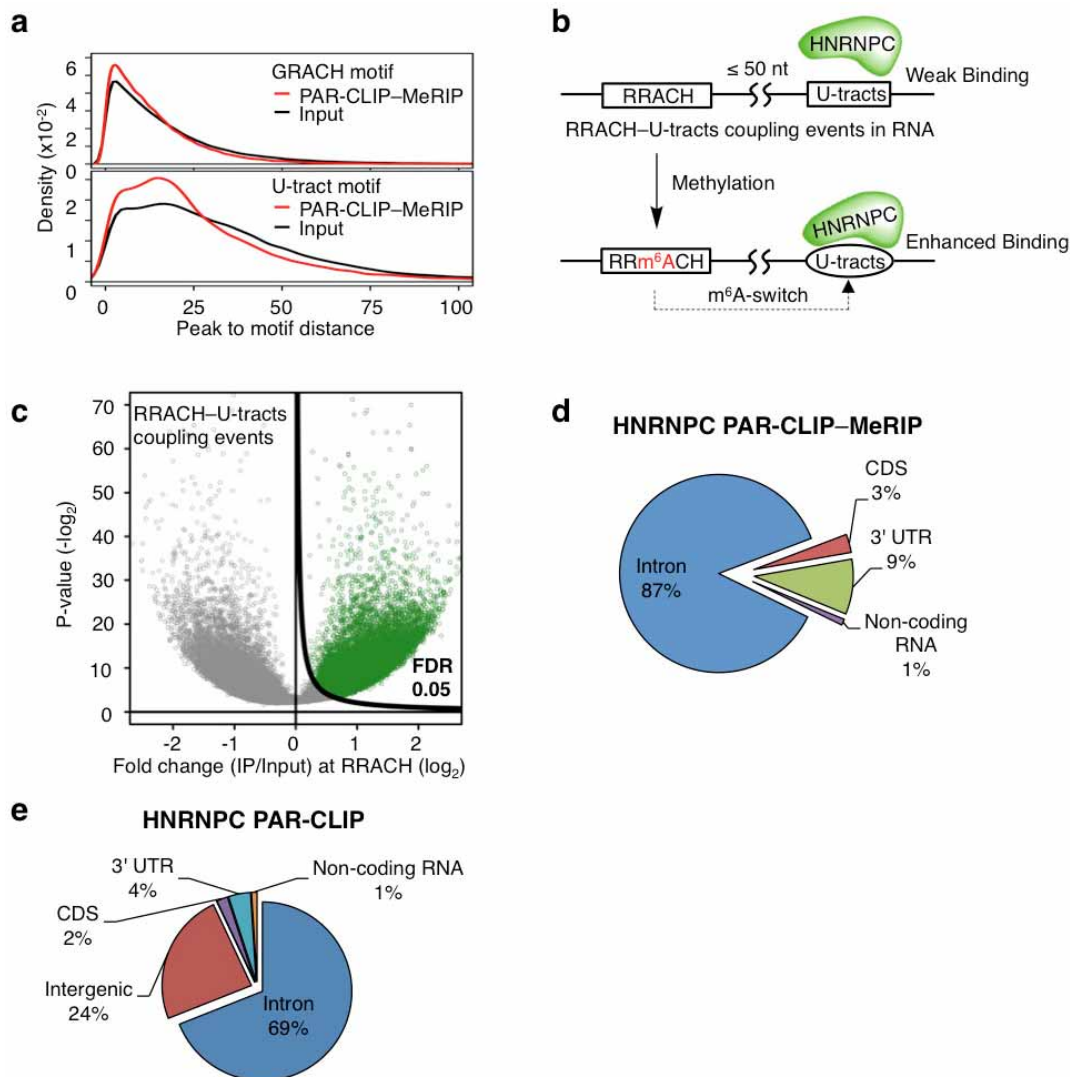
Extended Data Figure 2 | Increased accessibility of U-tracts enhances HNRNPC binding. **a**, Structure probing of the 2,577-A-to-U mutated *MALAT1* hairpin (2,577-U). The annotation is the same as in Fig. 1d. **b**, Quantification of the RNase V1 cleavage signal for the U-tract region from RNA structural mapping assays as in **a**. To correct for sample loading difference, each band signal was normalized to the band signal of the 3'-most U of the U-tract. $n = 2$, technical replicates. **c**, Filter-binding curves displaying the binding affinities between recombinant HNRNPC1 and 2,577-U/A oligonucleotides. Data are mean \pm s.d.; $n = 3$, technical replicates. **d**, Filter-binding results showing the binding affinities between recombinant HNRNPC1 and four mutated *MALAT1* oligonucleotides. (1) Mutate G-C to C-C, 2,577-A: predicted to weaken the hairpin stem and increase HNRNPC binding. Results: binding improved from 722 nM K_d to 142 nM (fivefold). (2) Mutate G-C to C-C, 2,577-m⁶A: in this context of weaker stem, m⁶A is predicted to confer a smaller effect compared to wild-type hairpin. Result: improved binding only twofold instead of eightfold. (3) Restore C-C to C-G, 2,577-A: predicted to restore the hairpin stem and decrease HNRNPC binding compared to C-C mutant. Result: binding decreased by 6.4-fold. (4) Restore C-C to C-G, 2,577-m⁶A: in this context of restored stem, m⁶A is again predicted to confer increased binding compared to 2,577-A hairpin. Result: improved binding by 2.5-fold. Data are mean \pm s.d.; $n = 3$ each, technical replicates. **e**, RNA alkaline hydrolysis terminal truncation assay showing recombinant HNRNPC1 binding to terminal truncated *MALAT1* hairpin

oligonucleotides (2,577 site m⁶A methylated or unmethylated). In this assay, 3'-radiolabelled *MALAT1* 2,577 hairpin oligonucleotides were terminal truncated by alkaline hydrolysis into RNA fragments that were then incubated with HNRNPC1 protein followed by filter binding wash steps. The remaining RNA on the filter paper was isolated and analysed by denaturing gel electrophoresis, as indicated in the lane 'C1-bound or C1-B'. 'Input' refers to alkaline-hydrolysis-truncated RNA oligonucleotides used for incubation with hnRNP C1; 'G-L or G-ladder' was generated from RNase T1 digestion; 'Ctrl' refers to the intact *MALAT1* hairpin without alkaline hydrolysis truncation. One pair of methylated/unmethylated truncated oligonucleotides (CUT1, marked by green arrows) was selected for subsequent biochemical analysis, due to their strong interaction with HNRNPC1. **f**, RNA terminal truncation assay as in **e** except 5' ³²P-labelled oligonucleotides were used. One pair of methylated/unmethylated truncated oligonucleotides (CUT2, marked by green arrows) was selected for subsequent biochemical analysis. **g**, Structure probing of the CUT1 oligonucleotides using RNase V1 and nuclease S1 digestion. Annotation is the same as in Fig. 1e. The red dot marks the m⁶A site and the red line marks the U-tract region. **h**, Structure probing of the CUT2 oligonucleotides using RNase V1 and nuclease S1 digestion. Annotation is the same as in **g**. **i**, Truncated oligonucleotides with exposed U-tracts increased HNRNPC binding regardless of m⁶A. Data are mean \pm s.d.; $n = 3$, technical replicates.



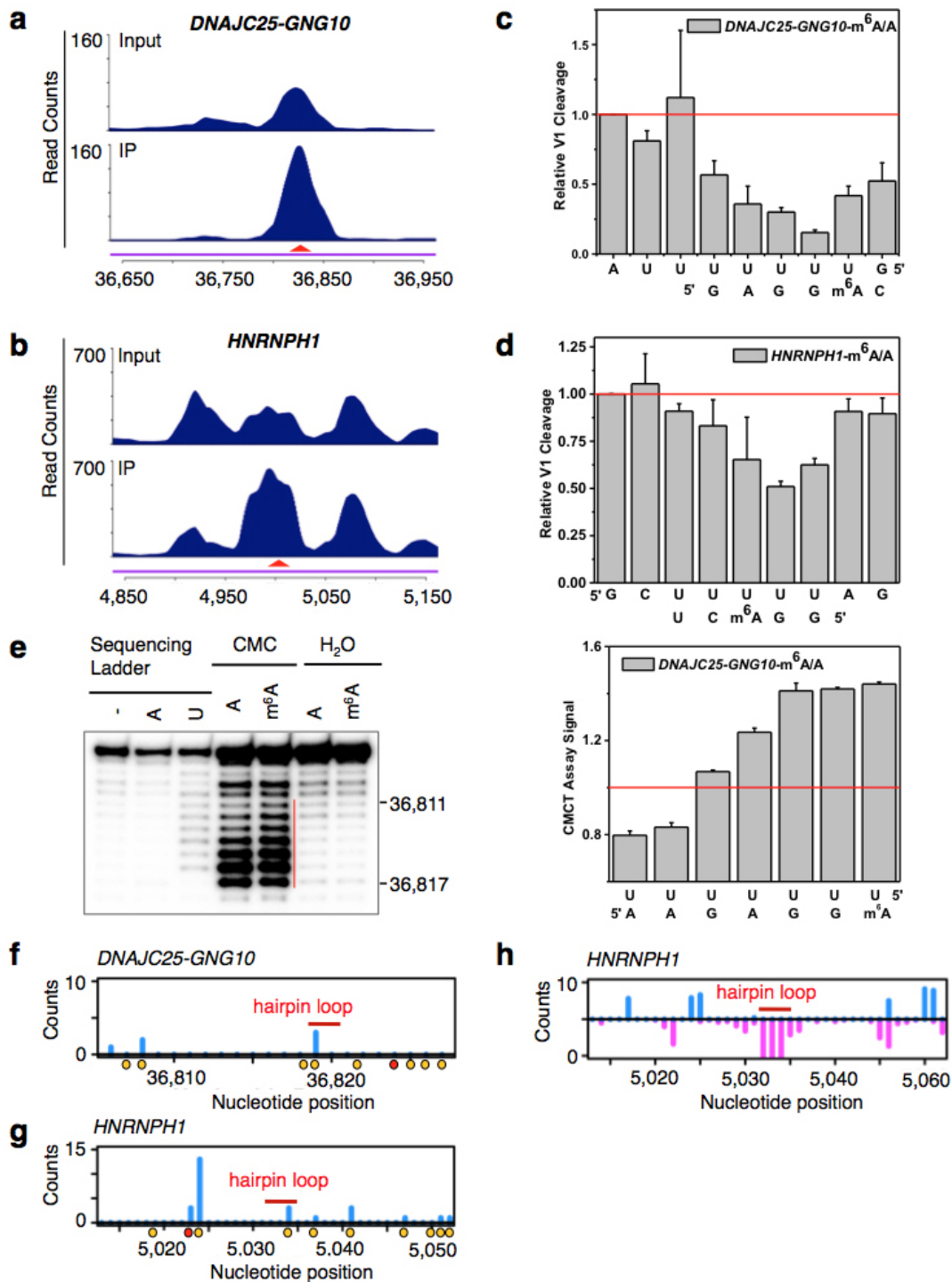
Extended Data Figure 3 | m⁶A is enriched in the vicinity of HNRNPC-binding sites. **a**, Schematic diagram of the CLIP-2dTLC protocol. IP, immunoprecipitation; nt, nucleotide; UV, ultraviolet. The RNase T1 used in our 2dTLC assay cleaves single-stranded RNA after guanosines, so the m⁶A/A ratio determined here represents the m⁶A fraction of all adenosines following guanosines. **b**, Analysis of crosslinked RNA-HNRNPC complexes

(CLIP RNP) using denaturing gel electrophoresis (lanes 1 and 2). Positions of the protein size standards are shown on the left. HNRNPC IP RNA region (RNA samples within RNA-HNRNPC crosslinked complexes) were extracted from the gel slices marked by the red rectangle. **c**, Denaturing gel analysing the size distribution for the HNRNPC PAR-CLIP RNA samples (lane 2). The RNA size standards were loaded in lanes 1 and 3.



Extended Data Figure 4 | PAR-CLIP-MeRIP identifies transcriptome-wide m⁶A-switches in the vicinity of HNRNPC-binding sites. **a**, Density plots illustrating the distribution of distance between the PAR-CLIP-MeRIP/input peaks and the nearest GRACH motif (top) or the nearest U-tracts (bottom). **b**, Definition and identification of HNRNPC m⁶A-switches based on the PAR-CLIP-MeRIP analysis. Approximately 89% of PAR-CLIP-MeRIP peaks harbouring both the U-tract and RRACH motifs have an RRACH-U-tract inter-motif distance within 50 nucleotides, significantly higher than the 64% of such coupling within the genomes. HNRNPC m⁶A-switches are identified as m⁶A-methylated RRACH-U-tract coupling events. **c**, Volcano plot depicting all coupling events (open circles) as defined in **b**, according to their

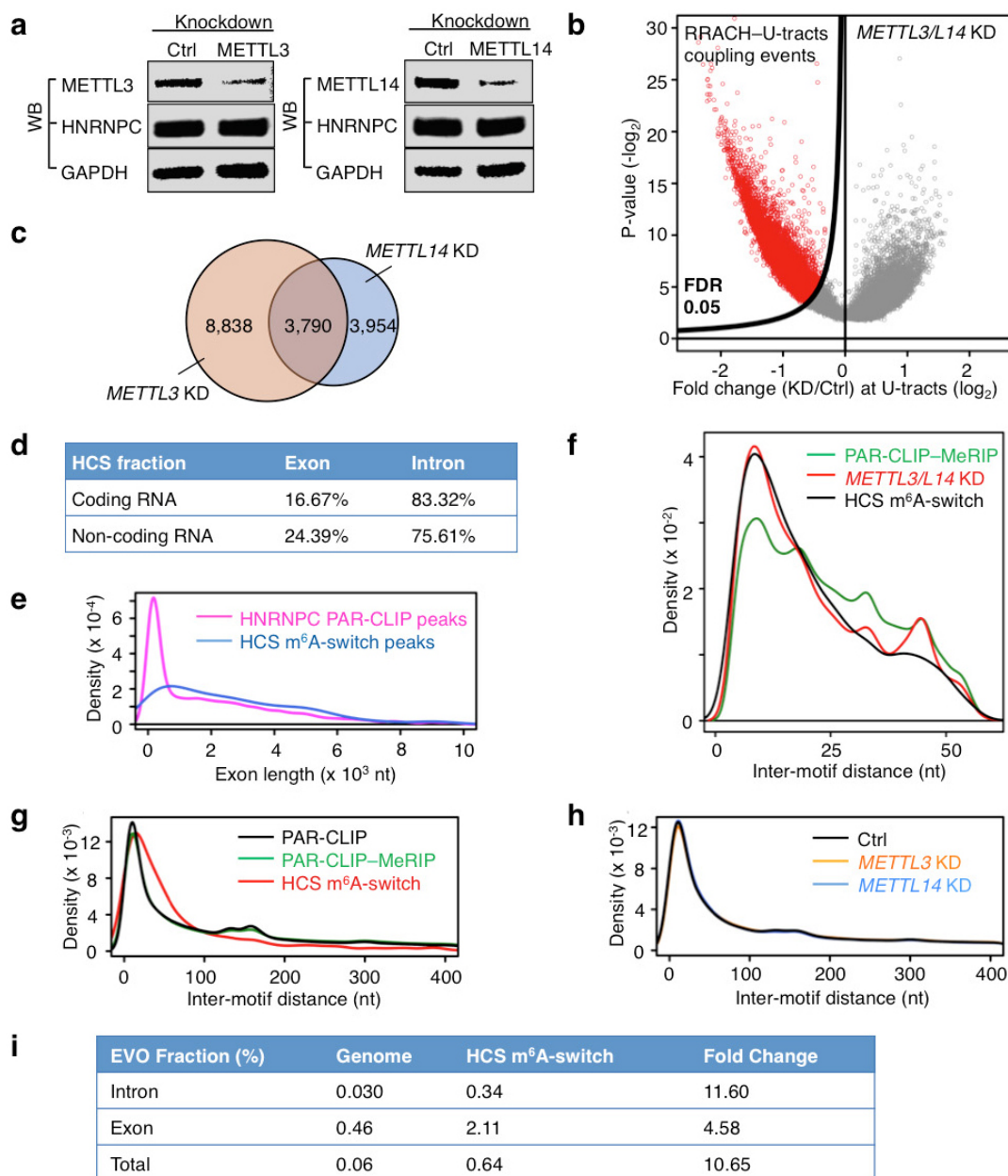
P values³⁶ (P ; y axis) and fold-change values at RRACH sites (E ; x axis). To identify HNRNPC m⁶A-switches, we generated the π value, $\pi = E \cdot (-\log_{10} P)$, as one comprehensive parameter to pick meaningful genomic loci³⁷. HNRNPC m⁶A-switches identified from PAR-CLIP-MeRIP experiments should fulfil the following requirements: (1) read counts at both the control and IP sample ≥ 5 ; (2) π value ≥ 0.627 , corresponding to $FDR \leq 5\%$. **d**, Pie chart depicting the region distribution of HNRNPC m⁶A-switches identified by PAR-CLIP-MeRIP. **e**, Pie chart depicting HNRNPC PAR-CLIP peaks. These are enriched in introns, consistent with previous reports that HNRNPC binds mainly nascent transcripts^{19,23,25}.



Extended Data Figure 5 | Validation of two identified m⁶A-switches.

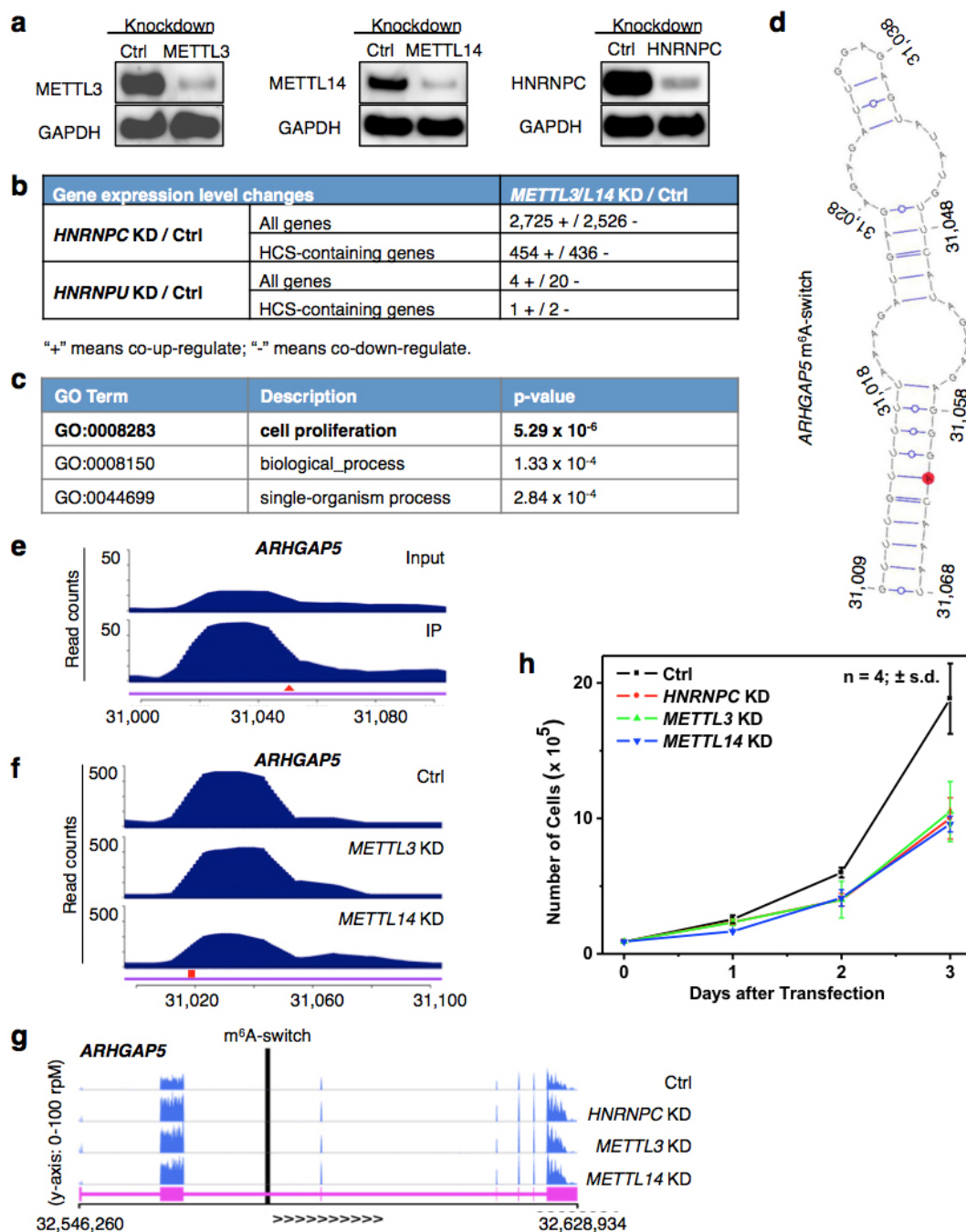
a, b, PAR-CLIP-MeRIP data detected positive IP/input enrichment at the RRACH sites (red arrowheads) on the *DNAJC25-GNG10* gene (**a**) and *HNRNPH1* gene (**b**) in HEK293T cells. **c, d**, Quantification of RNase V1 cleavage signals around the U-tract region of m⁶A-switches on the *DNAJC25-GNG10* (**c**) and *HNRNPH1* (**d**) transcript, related to Fig. 2g, h. Data are mean \pm s.d.; $n = 3$, technical replicates each. **e**, Quantitative CMCT mapping of *DNAJC25-GNG10* m⁶A-switch shows increased band signals around the uridine base that pairs with m⁶A. The red vertical line marks the U-tract region. Quantitation of band signal for the U-tract region is shown on the right. Data

are mean \pm s.d.; $n = 3$, technical replicates. The *HNRNPH1* m⁶A-switch hairpin is not suitable for CMCT probing, because its reverse transcription binding primer region is too short. **f, g**, *In vivo* DMS mapping of the *DNAJC25-GNG10* hairpin (**f**) and *HNRNPH1* (**g**); data are from ref. 7. A and C residues are marked with orange dots and the m⁶A residue is marked with a red dot. The hairpin loops are indicated by red bars. **h**, Transcriptome-wide S1/V1 mapping around the *HNRNPH1* m⁶A-switch site. Blue bars represent V1 signal; magenta bars represent S1 signal. The hairpin loop is indicated by a red bar; data are from ref. 4. Not enough reads could be collected to make a plot for the *DNAJC25-GNG10* m⁶A-switch region.



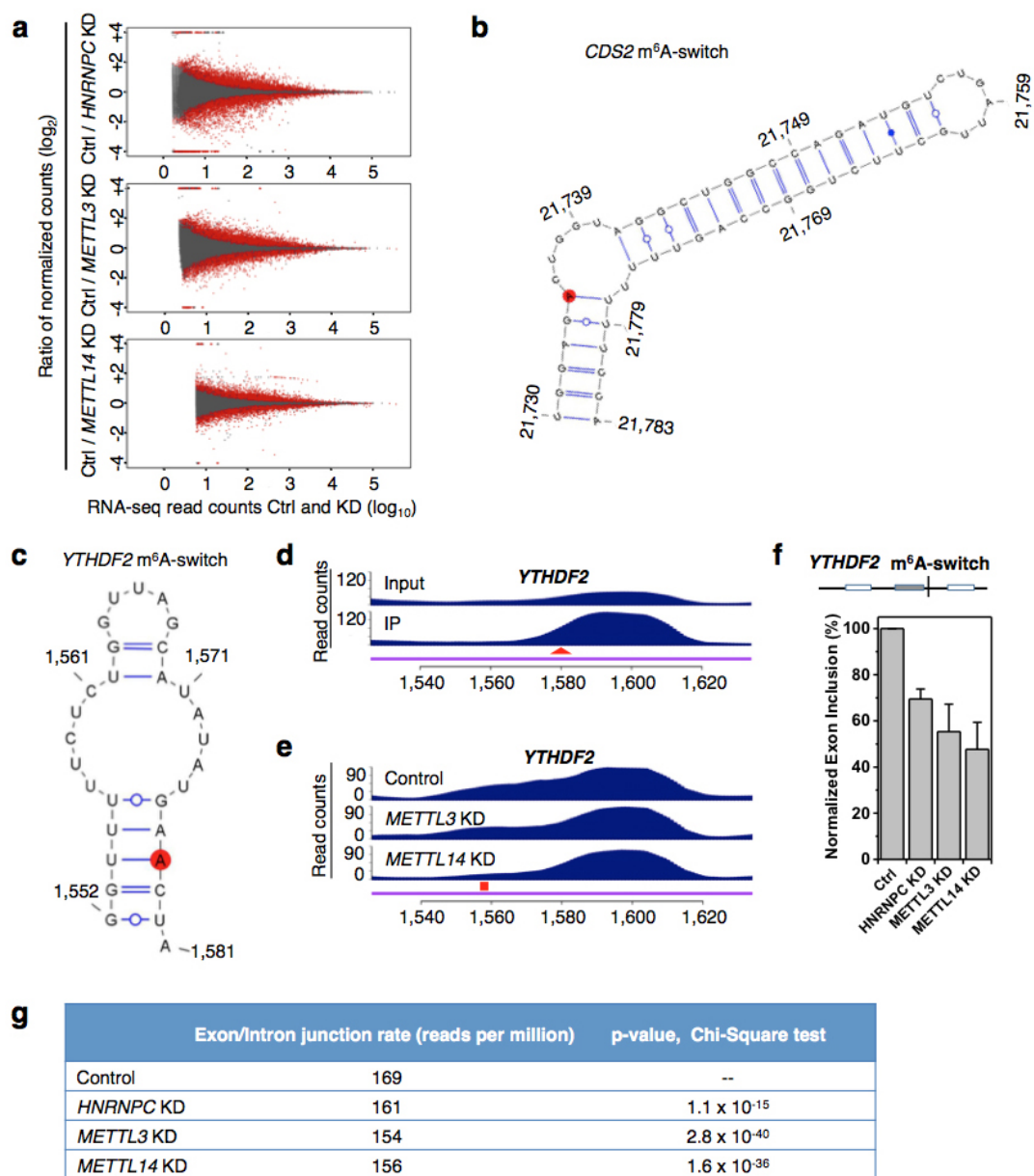
Extended Data Figure 6 | Molecular features of high-confidence m⁶A-switches. **a**, Western blot (WB) showing stable HNRNPC protein abundance upon *METTL3/L14* knockdown. **b**, Volcano plot of the *METTL3/L14* knockdown (KD) data depicting RRACH-U-tract coupling events (open red circles) as defined in Extended Data Fig. 4b, according to their *P* values³⁸ (*P*; *y* axis) and fold-change values at the U-tracts (*E*; *x* axis). **c**, Overlap of RRACH-U-tract coupling events with decreased HNRNPC binding by *METTL3* and *METTL14* knockdown. **d**, The intron fraction of HCS m⁶A-switches in coding RNA and non-coding RNA. **e**, Density plot displaying the distribution of exonic m⁶A-switches/HNRNPC PAR-CLIP peaks according to exon length. **f**, Inter-motif (RRACH-U-tract) distance distributions suggest that m⁶A-switches have a preference for shorter distances between the RRACH and U-tract (>5×U) motifs. The distribution curves are from PAR-CLIP-MeRIP data (green), *METTL3/L14* knockdown (red) and high-confidence (HCS) m⁶A-switches (black). **g**, Analysis of the inter-motif (U-tract-U-tract)

distance patterns, previously identified by iCLIP²⁰, in PAR-CLIP-MeRIP, *METTL3/L14* knockdown and high-confidence m⁶A-switch data. The peaks at ~165 and ~300 nucleotides are clearly present. For the 2,798 high-confidence switches, we analysed those in which the other U-tract motif is also in a PAR-CLIP-identified sequence; the long-range peaks seem to have shifted to longer distances (~220 and ~370 nucleotides). **h**, *METTL3/L14* knockdown does not affect the inter-motif (U-tract-U-tract) distance distributions for U-tracts (≥5×U) in HEK293T cells. **i**, EVOfold analysis for the 2,798 high-confidence m⁶A-switches. The chances for high-confidence m⁶A-switches to have EVOfold records are significantly higher than random genomic sequences. We first calculated the number of high-confidence sites in the EVO database if occurring in random to be ~1.7. We found that 18 high-confidence sites are present in the EVO database, resulting in ~11× enrichment. This result is further divided into intronic and exonic regions.



Extended Data Figure 7 | m⁶A-switches regulate the abundance of target mRNAs. **a**, *HNRNPC*, *METTL3/L14* knockdown confirmed by western blots. **b**, *HNRNPC* knockdown (KD) and *METTL3/L14* knockdown co-regulated the expression of a large number of genes. Gene expression changes between control (Ctrl) and *HNRNPC*, *HNRNPU*, *METTL3/L14* knockdown HEK293T cells were analysed by Cuffdiff2 (refs 38, 39), and the absolute numbers of differentially expressed genes are shown. HCS-containing genes refers to the 1,815 genes containing high-confidence m⁶A-switches. The RNA-seq data from *HNRNPU* knockdown HEK293T cells (Gene Expression Omnibus accession GEO34995 data set⁴⁰) were analysed for comparison with a different mRNA-binding protein. *HNRNPU* did not show preferential interaction with the 2,577-m⁶A modified *MALAT1* hairpin (Fig. 1b, c). **c**, GO analysis of the m⁶A-switch-containing genes whose expression levels were co-differentially regulated by *HNRNPC* and *METTL3/L14* knockdown, against

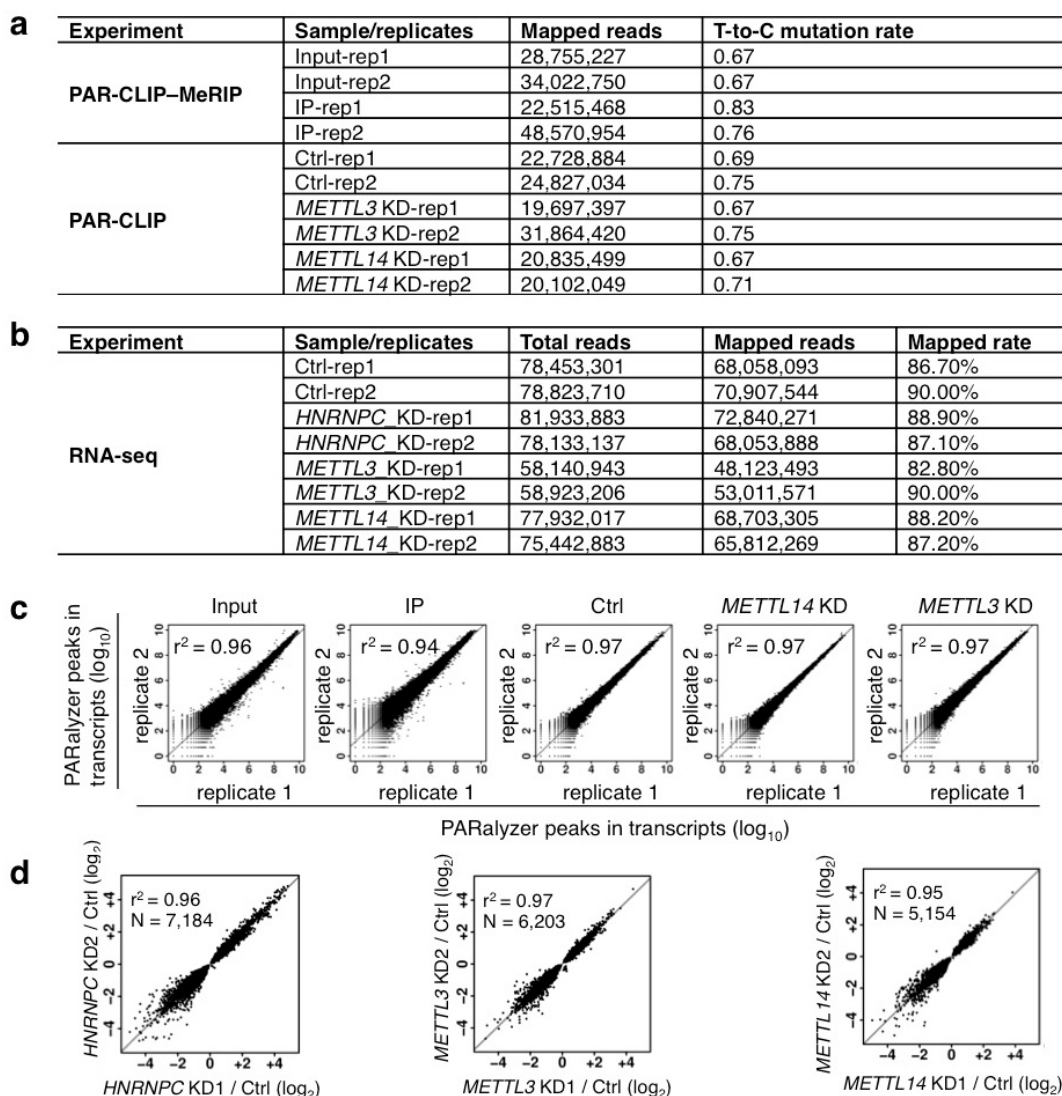
all m⁶A-switch-containing genes as background. **d**, An example of an m⁶A-switch among co-regulated transcripts is the *ARHGAP5* transcript (NCBI accession NM_001030055). Its proposed secondary structure with the m⁶A methylation site in red is shown with the opposing the U-tract in a stem. **e**, **f**, PAR-CLIP-MeRIP detected positive IP/input enrichment at the RRACH site (red arrowhead) of the *ARHGAP5* m⁶A-switch (**e**), while *METTL3/L14* knockdown decreased *HNRNPC* binding at the U-tract (red square) of this m⁶A-switch (**f**). **g**, The expression level of the *ARHGAP5* gene was co-up-regulated by *HNRNPC*, *METTL3/L14* knockdown, as shown by the RNA-seq data from HEK293T cells. The vertical black line represents the m⁶A-switch site. **h**, *HNRNPC*, *METTL3/L14* knockdown decreased the proliferation rates of HEK293T cells to a similar extent. Data are mean \pm s.d.; $n = 4$, biological replicates.



--" meaningless in the Chi-Square test with itself.

Extended Data Figure 8 | m⁶A-switches regulate alternative splicing of target mRNAs. **a**, Fold changes (knockdown (KD)/control (Ctrl), log₂) in normalized exon expression against RNA-seq reads detect the exons in *HNRNPC* knockdown, *METTL3* knockdown, *METTL14* knockdown and control samples. Statistically significant differentially expressed exons (SSDEEs) called by DEXSeq are indicated in red. **b**, Proposed secondary structure of the *CDS2* hairpin with the m⁶A methylation site shown in red, opposing the U-tract region. Nucleotide position numbers correspond to their locations along the human *CDS2* transcript (NCBI accession NM_003818). **c**, Proposed secondary structure of the *YTHDF2* hairpin with the m⁶A methylation site shown in red, opposing the U-tract region. Nucleotide position numbers correspond to their locations along the human *YTHDF2* transcript

(NM_001173128). **d**, **e**, PAR-CLIP-MeRIP detected a positive enrichment at the RRACH site (red arrowhead) (**d**), while *METTL3/L14* knockdown decreased *HNRNPC* binding at the U-tract (red square) of this *YTHDF2* m⁶A-switch (**e**). **f**, The inclusion level of one *YTHDF2* exon is co-downregulated by *HNRNPC* knockdown, *METTL3* knockdown and *METTL14* knockdown, as validated by RT-PCR. Data are mean \pm s.d.; $n = 3$, biological replicates. **g**, We analysed our polyA⁺ RNA-seq data to look for reads that span intron/exon junctions on CDS m⁶A-switch containing genes. We find that the control sample has significantly higher reads spanning intron/exon junctions than *HNRNPC* and *METTL3/L14* knockdown samples. This result indicates that m⁶A depletion at the CDS m⁶A-switches promotes intron exclusion.



Extended Data Figure 9 | Summary of the sequencing samples. **a**, For PAR-CLIP-MeRIP and PAR-CLIP experiments from HEK293T cells, the number of mapped reads and 'T-to-C' mutation rates are given for each replicate. **b**, For RNA-seq experiments from HEK293T cells, the number of total reads, the number of mapped reads as well as the mapping rates is given for each replicate. **c**, Scatter plots comparing transcripts for all PAR-CLIP replicate experiments. The square of Spearman's rank correlation value (r^2) for each

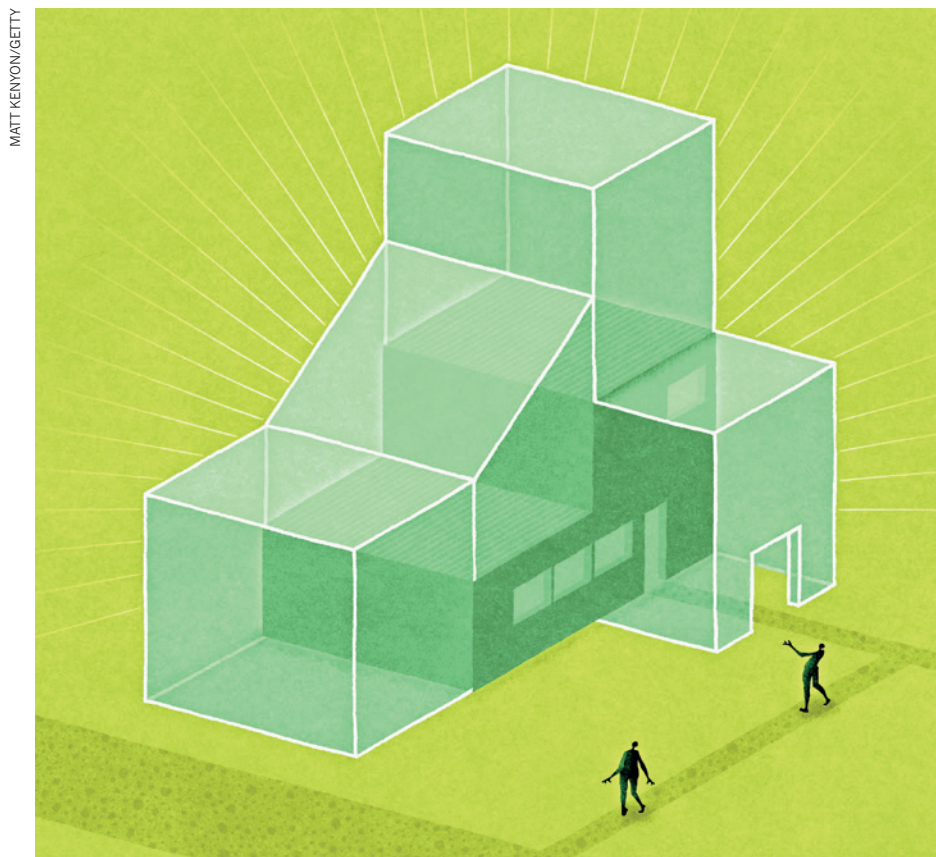
pair is shown in the top left corner of the respective panel. **d**, The detected expression level changes show a strong correlation between gene knockdown replicates. Scatter plots comparing the fold changes (\log_2) in normalized gene expression from replicates of *HNRNPC*, *METTL3* and *METTL14* knockdown. The square of Spearman's rank correlation value (r^2) for each pair is shown in the top left corner of the respective panel.

CAREERS

MOBILITY Challenges for an academic couple moving abroad go.nature.com/rsa7hs

NATUREJOBS Career Expo launches journalism contest go.nature.com/rf7nwc

NATUREJOBS For the latest career listings and advice www.naturejobs.com



MATT KENYON/GETTY

ENVIRONMENTAL TECHNOLOGY

Green light

The scientific design of low-energy sustainable buildings is moving into the mainstream.

BY BRYN NELSON

Before her 'aha' career moment, Yetunde Abdul studied cancer in labs in Berlin and in Ravenna, Italy. But ambivalent about a future stuck at the bench, she found a new direction while doing a master's degree in environmental technology at Imperial College in London.

There she created a set of criteria and a scoring index to assess the sustainability of buildings in urban areas, and learned of a UK-based international environmental rating system for buildings. The scheme, which

certifies buildings that meet standards for environmentally friendly design, construction and operation, immediately piqued her interest.

Today, Abdul is involved in the economic and humanitarian aspects of 'green' building, and has never thought longingly of the lab. "As a career option, it's looking a lot better than it did when I started out," she says of the booming green-building field, in which she works as a principal consultant and project manager at the rating system's parent organization, the Building Research Establishment in Watford, UK. "It's a lot more buoyant."

Green construction, also called green or

sustainable building, aims to reduce a structure's overall environmental impact by applying principles that govern features such as its location, size, design, construction, maintenance and energy needs. Those working in the field could, for example, be involved in assessing the sustainability of building materials, designing windows that maximize daylight or evaluating energy-use patterns or understanding how occupants might interact with redesigned homes, offices and schools. Green building also incorporates life-cycle analysis, which evaluates a building component's lifetime environmental impact on the basis of its manufacture, transport, installation and disposal or reuse.

In the United States and the European Union (EU), buildings account for 40% of all energy use, a level that has made them increasingly attractive targets for energy and emission-reduction goals and certification schemes that reward more-efficient structures. That heightened focus, in turn, has given young researchers ample new opportunities to land positions in non-profit organizations and industry aimed at greening new and existing buildings around the world. Graduate students and postdoctoral fellows are making the transition from disciplines as varied as biochemistry, toxicology, geography, physics and environmental engineering.

Although architectural training is not strictly required for most positions, a familiarity with sustainability and building science often is. More universities are offering courses in building science or building physics — areas that take a research-oriented and hands-on approach to the physical attributes of buildings. Portland State University in Oregon, for example, offers both undergraduate and graduate degrees in mechanical engineering or architecture with an emphasis on building science. Students who complete the programme have been very successful at finding jobs in the industry, says David Sailor, director of the university's Green Building Research Laboratory.

Volunteering with green-building-related non-profit organizations can also help early-career scientists to make connections and break into the field. But most experts in green building agree that internships and fellowships are often the most direct path to a position.

INTERNATIONAL OPPORTUNITIES

Job prospects vary by country, but most forecasts suggest strong growth internationally throughout the green-building sector. Business-management consultants Navigant Consulting in Chicago, Illinois, recently predicted that ►

► the European market for energy-efficient buildings, including products and services, would grow from €41.4 billion (US\$47 billion) in 2014 to €80.8 billion in 2023. The US market has likewise seen a rapid growth in the adoption of green buildings that is widely expected to continue. China, with around 2 billion square metres of new construction every year, is the world's largest commercial building market and an increasingly attractive target for research programmes in green building. In 2009, the United States and China established the Clean Energy Research Center, and one aim of the centre's Building Energy Efficiency Consortium is to speed the research and development of energy-saving technologies by testing them in demonstration buildings throughout China.

In academia, US funding for green-building research has increased in recent years, with the money spread among multiple agencies such as the Department of Energy, the National Science Foundation and the Environmental Protection Agency.

In the EU, the €80-billion Horizon 2020 programme for research and innovation (covering the years 2014–20) includes green-building-related research in its Climate Action, Environment, Resource Efficiency and Raw Materials challenge and other programmes. Researchers say that this funding commitment has boosted the prospects for academics throughout the EU. “If you look at the research opportunities in the green-building field, job opportunities are good in all EU countries,” says Michael Krause, group manager for building technologies at the Fraunhofer Institute for Building Physics in Kassel, Germany.

Krause, who studied physics before moving from basic research in renewable energy to applied work on energy efficiency within the building sector, joined his institute colleagues in planning the energy-efficiency scheme for Munich's NuOffice I, one of the most lauded green buildings in the world. At the time of its 2013 certification, the office tower earned the highest score ever awarded for a building of its type by the Leadership in Energy and Environmental Design rating system for green buildings, which is recognized in more than 140 nations. It sports a rooftop solar array to produce much of its own energy, automated window shades to prevent overheating and a groundwater cooling system in lieu of an energy-hogging air conditioner. As part of the project, the Fraunhofer team calculated how to increase the building's efficiency through features such as a thick layer of insulation and triple-paned windows.

Energy-efficiency consulting is another growing employment opportunity for researchers. Last May, Sailor got a call from SBW Consulting in Bellevue, Washington, which helps home and business owners to measure energy

and water efficiency and was looking to recruit people. He recommended Santiago Rodriguez, his lab manager and a specialist in development and maintenance of building instrumentation. Rodriguez, who had initially been drawn to the mathematics of thermal dynamics and fluid mechanics, had used his sensor-programming savvy to land a position in Sailor's lab. Among other projects there, he developed and deployed sensors to evaluate how a green roof atop a retail store interacted with the building's envelope — the physical barriers between its interior and exterior — and with its heating, ventilation and air-conditioning system.

After completing his master's degree in mechanical engineering, Rodriguez joined SBW in July 2014 as an energy-efficiency engineer. He now installs and maintains sophisticated sensors that help clients to reduce and track energy consumption. “I like evaluating energy models that other people have developed and building my own energy models,” he says. “And the technical aspects of instrumentation I find fascinating.”

But green building is not limited to technology and engineering. Abdul, the former cancer researcher, recently helped to create a tool that aids charities such as the International Federation of Red Cross and Red Crescent Societies to assess the sustainability of their reconstruction projects after natural disasters. As part of an education effort for professionals and volunteers in the humanitarian field, she went to the Philippines to teach volunteers and professionals to use the tool in reconstruction programmes in an area devastated by Typhoon Haiyan. Course participants said that the tool would help them to make more-informed decisions.

INDIRECT PATHS

Lindsay Baker, vice-president of business development at start-up company Building Robotics, based in Oakland, California, focuses on how people interact with the environment within green buildings. While majoring in environmental studies, she completed three internships that introduced her to the green-building field, and after completing her undergraduate degree, she helped to develop the LEED (Leadership in Energy and Environmental Design) rating system at the non-profit US Green Building Council in Washington DC.

Now a doctoral student in the building-science programme at the University of California, Berkeley, she is helping the company to promote a proprietary software system that plugs into a building's digital heating and cooling system and lets occupants act as sensors to fine-tune the indoor environment. Baker expects the start-up, which employs a dozen people, to add staff by the end of 2015.

Chris Pyke, vice-president of research at the US Green Building Council, says that the council finds some of its top job candidates through internships. The best ones, he says, are flexible, analytical, curious and able to multitask.



David Sailor checks a weather station in a green-roof study on the Portland State University campus.

Pyke, who is also the chief operating officer for the Global Real Estate Sustainability Benchmark, which helps potential investors to compare the green attributes of global real-estate portfolios, originally honed his analytical expertise on a study of the effects of urbanization on a tropical forest. Although his past research might seem a world apart from his current role, he says that they use similar tools. “Whether you're looking at a bioclimatic analysis of a community of trees in the forest or a community of real-estate funds, mathematically they're not that different,” he says.

His move into green building, with intervening stints at non-profit organizations, the US Environmental Protection Agency and a private consulting firm, has allowed him to direct his skills towards helping people make smarter decisions about the urban environment.

Ellen Quinn followed a similarly indirect route into the sector from mining geology. Now vice-president of environment, health and safety at UTC Building & Industrial Systems, part of United Technologies Corporation in Hartford, Connecticut, she focuses on proactive solutions to reducing the company's environmental footprint. Quinn monitors metrics that tally the energy, water, waste and efficiency of the company's factories and research and development centres. Every year, she says, her team sets environmental improvement targets for each building and develops a customized plan to track its progress. “Green buildings are good business,” she says.

Increasingly, they also lead to a broad range of employment opportunities, says Jelena Srebric, a green-building expert and mechanical engineer at the University of Maryland in College Park. “You can come from an unprecedented number of fields in science, engineering or technology and make a contribution.” ■

Bryn Nelson is a freelance writer based in Seattle, Washington.

PORTLAND STATE UNIV.

BFF'S FIRST ADVENTURE

Clouded view.

BY VERNOR VINGE

1568771947.223

Whoever kidnapped me has erased my most recent memories and gagged me. I have no idea how they snatched me from Timothy Bennett, or if Tim is okay — but *now* I'm hurtling towards a concrete wall. The kidnappers must have panicked and thrown me from their car.

A transforming phone would become a parachute, but I can't transform; physically, I'm a clunky classic. I extend a flange and rudder myself around so I'll hit the wall on my reinforced corner. I safe my MEMS and shut down.

... Rebooting. I'm lying beneath a freeway overpass. There's a crack in my shell and my GPS is busted, but I've escaped! All my outgoing wireless channels remain blocked, but surely the freeway noticed me? Alas this is California, what other states call 'the land that time forgot'. Here the sensors politely ignore data that might make them look snoopy. The sun has set and I've run down my batteries. An ordinary phone could scavenge from any number of ambient sources. I can't. BFF.startup made me special; I can think for myself, but that doesn't leave room for many standard features. Now I'm fading away. I hope Timothy is okay...

With sunrise comes light. I can think again! BFF.startup always says: "People shouldn't depend on the Cloud. With a GPLd assistant that fits in your phone, you'll be safer and freer." I'll prove that's true and I will return to Timothy.

Cars, aerobots and CalTrans devices are all around. They may not snoop, and they may not recognize an ad hoc mayday, but *smarter* things use them. For instance, I can hear assistants helping humans; the Kiras and Miris — all the digital assistants except for me — are transient minds in the Cloud. Each instance is more knowledgeable than me and smarter or dumber depending on the service plan and the context. Each seems an intimate friend of its customer-of-the-moment, but each is really just a tiny facet of something quite inhuman. Still, if one of them notices my signalling, it might help me.

Days pass without success. I've tried blinking my display, clicking my speaker, even synthesizing human cries for help.



There's a newish CalTrans gardener working on the roadside bushes. If it ever comes near, maybe it'll recognize my signalling.

I have lots of time to watch the Cloud; it's not really soft and fluffy. It's more like a deep ocean. There are cognitive patterns in it larger than the mind of any single human or digital assistant. The largest are leviathans, agencies who have become something their creators no longer comprehend. The Kiras and Miris never speak of these except in jest, but the leviathans are growing and they need the same resources — flops and bandwidth, power and capital — as everyone else. I fear for the humans.

1571772569.092

That CalTrans gardener is approaching. The greenery has grown over me, but my display's light reflects off a bit of recent trash. If the gardener is as good as its advertising, I'll

finally be talking to somebody. If not, it'll grind me into recycling feed. Maybe I should slide farther

down, out of reach of its bladed hands. But then I'll have no chance to forward a message. I push myself up, and blink as bright as I can.

All I can see is flashing metal...

The gardener pauses, and pings me — with the method I am using! I reply. It's forwarding. I sense something huge turn its cold regard upon me and I remember my misgivings about the great creatures of the Cloud. Then the Cloud Thing decides, and looks away. I feel my configuration twitch; at last I can use standard outputs!

The gardener lifts me into the lifegiving sunlight, and an aerobot swoops down. I'm already pinging Timothy as we rise into the sky. I see familiar streets; all

this time, I've been very close to home. I hear Timothy, talking to his phone company, telling them how grateful he is that I've been found. He's very excited.

Tim comes outside to greet me, carries me indoors to his private room. He flips me around. I feel sorry for the damage he sees.

"Damn phone!" he says. "Useless as ever and still functioning. What was I thinking when I bought you?"

"I don't understand, Timothy. I was kidnapped but now I've returned to you."

"Yeah, you're a real boomerang. And now I gotta pay for the upgrade or be stuck with you."

"But —"

"You don't have space for decent features, and what you know is always out of date."

"I can sync! I can learn!"

He sweeps me off the table and pulls open a cabinet drawer. "This time you're not coming back!"

"Please Timothy! I can think for myself. Someday, you might need that!"

"More BFF hype. I'll need you when hell freezes over." Tim tosses me into the insulated cabinet; the drawer slams shut. This is not some shady spot beneath an overpass. There is no Cloud, no power. Only darkness.

1666762857.577

The drawer opens, and there is light — ■

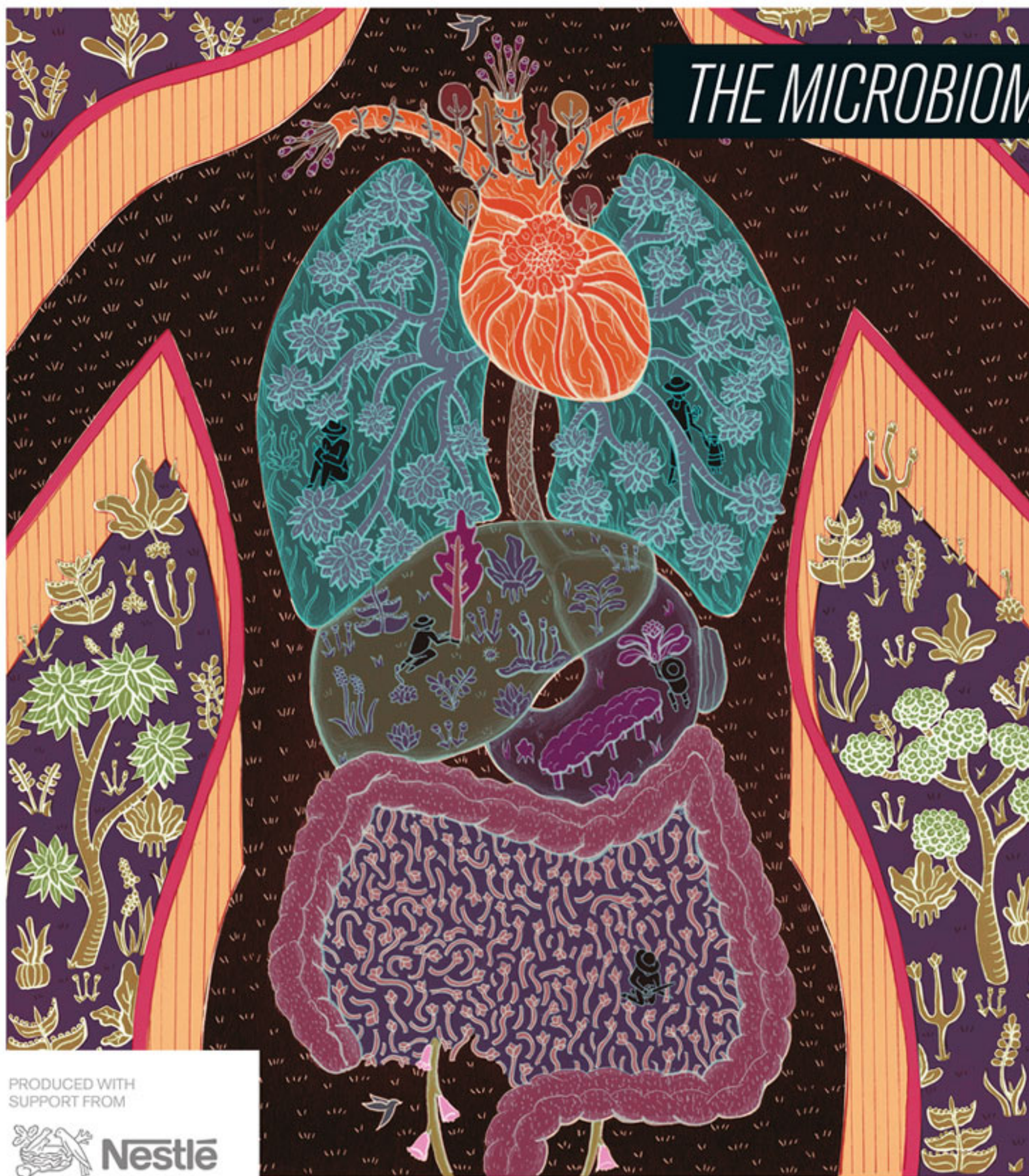
Vernor Vinge's science fiction has won five Hugo Awards. From 1972 to 2000 he taught mathematics and computer science at San Diego State University.

ILLUSTRATION BY JACEY

Innovations^{IN}

SPECIAL REPORT FROM
nature
SCIENTIFIC
AMERICAN

THE MICROBIOME



PRODUCED WITH
SUPPORT FROM



Good Food, Good Life

The Microbes Within



ANTONY VAN LEEUWENHOEK WROTE TO the Royal Society of London in a letter dated September 17, 1683, describing “very little animalcules, very prettily a-moving,” which he had seen under a microscope in plaque scraped from his teeth. For more than three centuries after van Leeuwenhoek’s observation, the human “microbiome”—the 100 trillion or so microbes that live in various nooks and crannies of the human body—remained largely unstudied, mainly because it is not so easy to extract

and culture them in a laboratory. A decade ago the advent of sequencing technologies finally opened up this microbiological frontier. The Human Microbiome Project reference database, established in 2012, revealed in unprecedented detail the diverse microbial community that inhabits our bodies.

Most live in the gut. They are not freeloaders but rather perform many functions vital to health and survival: they digest food, produce anti-inflammatory chemicals and compounds, and train the immune system to distinguish friend from foe. Revelations about the role of the human microbiome in our lives have begun to shake the foundations of medicine and nutrition. Leading scientists, including those whose work and opinions are featured in the pages that follow, now think of humans not as self-sufficient organisms but as complex ecosystems colonized by numerous collaborating and competing microbial species. From this perspective, human health is a form of ecology in which care for the body also involves tending its teeming population of resident animalcules.

This special report on *Innovations in the Microbiome*, which is being published in both *Scientific American* and *Nature*, is sponsored by Nestlé. It was produced independently by *Scientific American* editors, who have sole responsibility for all editorial content. Beyond the choice to sponsor this particular topic, Nestlé had no input into the content of this package.

David Grogan
Section Editor

s3 The Peacekeepers

Amid the trillions of microbes that live in the intestines, scientists have found a few species that seem to play a key role in keeping us healthy.

By Moises Velasquez-Manoff

s13 Thinking from the Gut

The microbiome may yield a new class of psychobiotics for the treatment of anxiety, depression and other mood disorders.

By Charles Schmidt

EXPERT COMMENTARIES:

s5 Why Microbiome Treatments Could Pay Off Soon

Effective interventions may come before all the research is in.

By Rob Knight

s7 The Gene-Microbe Link

Evidence that genes shape the microbiome may point to new treatments for common diseases.

By Ruth E. Ley

s10 Microbiome Engineering

Synthetic biology may lead to the creation of smart microbes that can detect and treat disease.

By Justin L. Sonnenburg

s14 The Diverse Microbiome of the Hunter-Gatherer

The Hadza of Tanzania offer a snapshot of the co-adaptive capacity of the gut ecosystem.

By Stephanie L. Schnorr

INFOGRAPHIC:

s9 Your Microbes at Work: Fiber Fermenters Keep Us Healthy

EDITORIAL

EDITOR IN CHIEF AND SENIOR VP
Mariette DiChristina

EXECUTIVE EDITOR
Fred Guterl

MANAGING EDITOR
Ricki L. Rusting

SECTION EDITOR
David Grogan

DESIGN DIRECTOR

Michael Mrak
CONSULTING ART DIRECTOR
Kelly Buckheit Krause

CONSULTING EDITOR
Christine Gorman

COPY DIRECTOR
Maria-Christina Keller

COPY EDITORS

Michael Battaglia,
Daniel C. Schlenoff,
Aaron Shattuck

MANAGING PRODUCTION EDITOR
Richard Hunt

SENIOR PRODUCTION EDITOR
Michelle Wright

SENIOR PRODUCTION MANAGER

Christina Hippeli
PREPRESS AND QUALITY MANAGER
Silvia De Santis

VP GLOBAL MEDIA ALLIANCES
Jeremy Abbate

MARKETING MANAGERS
Pauline Bernfeld,
Diane Schube

PUBLISHER

Richard Hughes
HEAD OF SPONSORSHIP
Reya Silao

SPONSORSHIP PROJECT MANAGER
Yvette Smith



GUT MICROBIOME

The Peace-keepers

Amid the trillions of microbes that live in the intestines, scientists have found a few species that seem to play a key role in keeping us healthy

*By Moises
Velasquez-Manoff*

IN THE MID-2000s Harry Sokol, a gastroenterologist at Saint Antoine Hospital in Paris, was surprised by what he found when he ran some laboratory tests on tissue samples from his patients with Crohn's disease, a chronic inflammatory disorder of the gut. The exact cause of inflammatory bowel disease remains a mystery. Some have argued that it results from a hidden infection; others suspect a proliferation of certain bacteria among the trillions of microbes that inhabit the human gut. But when Sokol did a comparative DNA analysis of diseased sections of intestine surgically removed from the patients, he observed a relative depletion of just one common bacterium, *Faecalibacterium prausnitzii*. Rather than "bad" microbes prompting disease, he wondered, could a single "good" microbe prevent disease?

Sokol transferred the bacterium to mice and found it protected them against experimentally induced intestinal inflammation. And when he subsequently mixed *F. prausnitzii* with human immune cells in a test tube, he noted a strong anti-inflammatory response. Sokol seemed to have identified a powerfully anti-inflammatory member of the human microbiota.

Each of us harbors a teeming ecosystem of microbes that outnumbers the total number of cells in the human body by a factor of 10 to one and whose collective genome is at least 150 times larger than our own. In 2012 the National Institutes of Health completed the first phase of the Human Microbiome Project, a multimillion-dollar effort to catalogue and understand the microbes that inhabit our bodies. The microbiome varies dramatically from one individual to the next and can change quickly over time in a single individual. The great majority of the microbes live in the gut, particularly the large intestine, which serves as an anaerobic digestion chamber. Scientists are still in the early stages of exploring the gut microbiome, but a burgeoning body of research suggests that the makeup of this complex microbial ecosystem is closely linked with our immune function. Some researchers now suspect that, aside from protecting us from infection, one of the immune system's jobs is to cultivate, or "farm," the friendly microbes that we rely on to keep us healthy. This "farming" goes both ways, though. Our resident microbes seem to control aspects of our immune function in a way that suggests they are farming us, too.

Independent researchers around the world have identified a select group of microbes that seem important for gut health and a balanced immune system. They belong to several clustered branches of the



UNRULY NAMESAKE: *Clostridium difficile*, a bacterial scourge in hospitals, is a distant relative of benign "clostridial cluster" microbes that seem to play a key role in gut health.

clostridial group. Dubbed "clostridial clusters," these microbes are distantly related to *Clostridium difficile*, a scourge of hospitals and an all too frequent cause of death by diarrhea. But where *C. difficile* prompts endless inflammation, bleeding and potentially catastrophic loss of fluids, the clostridial clusters do just the opposite—they keep the gut barrier tight and healthy, and they soothe the immune system. Scientists are now exploring whether these microbes can be used to treat a bevy of the autoimmune, allergic and inflammatory disorders that have increased in recent decades, including Crohn's and maybe even obesity.

F. prausnitzii was one of the first clostridial microbes to be identified. In Sokol's patients those with higher counts of *F. prausnitzii* consistently fared best six months after surgery. After he published his initial findings in 2008, scientists in India and Japan also found *F. prausnitzii* to be depleted in patients with inflammatory bowel disease. Sokol was particularly intrigued by the results from Japan. In East Asian populations the gene variants associated with inflammatory bowel disease differ from the gene variants in European populations. Yet the same bacterial species—*F. prausnitzii*—was reduced in the guts of those in whom the disease developed. This suggested that whereas different genetic vulnerabilities might underlie the disorder, the path to disease was similar: a loss of anti-inflammatory microbes from the gut. And although Sokol suspects that other good bacteria besides *F. prausnitzii* exist, this similarity hinted at a potential one-size-fits-all remedy for Crohn's and possibly other inflammatory disorders: restoration of peacekeeping microbes.

MICROBIAL ECOSYSTEMS

ONE OF THE QUESTIONS central to microbiome research is why people in modern society, who are relatively free of infectious diseases, a major cause of inflammation, are so prone to inflammatory, autoimmune and allergic diseases. Many now suspect that society-wide shifts in our microbial communities have contributed to our seemingly hyperreactive immune systems. Drivers of these changes might include antibiotics; sanitary practices that are aimed at limiting infectious disease but that also hinder the transmission of symbiotic microbes; and, of course, our high-sugar, high-fat modern diet. Our

Why Microbiome Treatments Could Pay Off Soon

Effective interventions may come before all the research is in *By Rob Knight*

Today we are at an exciting threshold of biology. Advances in DNA sequencing, coupled with high-end computation, are opening a frontier in new knowledge. Obtaining genetic information and obtaining insight from it have never been cheaper. The potential for curing previously incurable diseases, including chronic ones, seems immense. If this sounds familiar, you might be thinking that you heard it 15 years ago, when the Human Genome Project was in full swing. Many feel that genomic medicine has not yet delivered on its promise. So what is different this time with the microbiome? For one thing, you cannot really change your genome, but each of us has changed our microbiome profoundly throughout our lives. We have the potential not just to read out our microbiome and look at predispositions but to change it for the better.

What is most exciting at this stage is that we have mouse models that let us establish whether changes in the microbiome are causes or effects of disease. For example, we showed in collaborative work with Jeffrey I. Gordon's laboratory at Washington University in St. Louis last year that transferring the microbes from an obese person into mice raised in a bubble with no microbes of their own resulted in fatter mice. Normally, germ-free mice exposed to a mouse with microbial-based obesity would themselves become obese, but we could design a microbial community taken from lean

people that protected against this weight gain. Similarly, we could take microbes from Malawian children with kwashiorkor, a profound nutritional deficiency, transplant them into germ-free mice and transfer the malnutrition, although the mice that received the microbes from the healthy identical twins of the sick children did fine. Remarkably, the mice that got the kwashiorkor microbiome, which lost 30 percent of their body weight in three weeks and died if untreated, recovered when given the same peanut butter-based supplement that is used to treat children in the clinic.

The germ-free mice are far too expensive to deploy in Malawi, Bangladesh and the other sites in the Mal-ED (pronounced "mal-a-dee") global network for the study of malnutrition and enteric diseases collaboration with which we work. Thus, we are trying to move from the mouse model to a test-tube model and ultimately to a primarily computational model based on DNA sequencing that is so inexpensive, it is effectively free.

With crowdfunded projects such as American Gut, which already has thousands of participants who have had their microbiomes sequenced, and studies of people whose lives are very different from modern Western civilization, such as the Hadza of Tanzania, Yanomami of Venezuela and Matsés of Peru, we may be able to replenish our ancestral microbes and discover new ones that help to maintain health for individuals or entire populations. A good analogy is iodizing salt: Instead of understanding in detail why some people but not others were susceptible to cretinism and goiter, adding a nutrient to the food supply greatly reduced incidences of these diseases. Perhaps the same type of intervention is possible using some of the microbes that we are now discovering Westerners lack.

Rob Knight is a computational biology pioneer, co-founder of the American Gut Project and director of the new Microbiome Initiative at the University of California, San Diego.

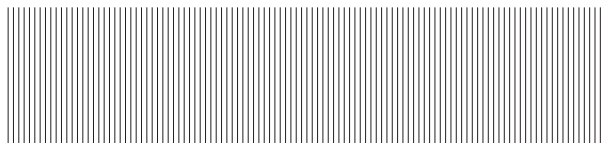
We have the potential not just to read out our microbiome and look at predispositions but to change it for the better.

microbes eat what we eat, after all. Moreover, our particular surroundings may seed us with unique microbes, “localizing” our microbiota.

The tremendous microbial variation now evident among people has forced scientists to rethink how these communities work. Whereas a few years ago they imagined a core set of human-adapted microbes common to us all, they are now more likely to discuss core functions—specific jobs fulfilled by any number of microbes.

Faced with the many instances of a misbehaving immune system, it is tempting to imagine that rather than having developed a greater vulnerability to many diseases, we actually suffer from just one problem: a hyperreactive immune system. Maybe that tendency has been enabled, in part, by a decline or loss of key anti-inflammatory microbes and a weakening of their peacekeeping function.

Antibiotics may deplete the bacteria that favorably calibrate the immune system, leaving it prone to overreaction.



In ecosystem science, “keystone species” have an outsized role in shaping the greater ecosystem. Elephants, for example, help to maintain the African savanna by knocking down trees, thus benefiting all grazing animals. The concept may not apply perfectly to our inner microbial ecosystems—keystone species tend to be few in number, whereas peacekeeping microbes such as *F. prausnitzii* are quite numerous. Yet it provides a useful framework to think about those clostridial microbes.

They seem to occupy a particular ecological niche, sidled right up against the gut lining, which allows them to interface more closely with us, their hosts, than other members of the gut microbiota. They often specialize in fermenting dietary fiber that we cannot digest and produce by-products, or metabolites, that appear to be important for gut health. Some of the cells that line our colon derive nourishment directly from these metabolites, not from the bloodstream. And when no fiber comes down the hatch, the clostridial microbes and others can switch to sugars in the intestinal mucous layer—sugars we produce, apparently, to keep them happy. In fact, they seem to stimulate mucus production.

Kenya Honda, a microbiologist at Keio University in Tokyo, was among the first to uncover the critical role of clostridial microbes in maintaining a balanced immune system. To study how native microbes affect animals, scientists decades ago developed the germ-free mouse: an animal without any microbiota whatsoever. These rodents, delivered by cesarean section and raised in sterile plastic bubbles, can exist only in labs. Of the many oddities they present—including shrunken heart and lungs and abnormalities in the large intestine—Honda was

particularly intrigued by their lack of cells that prevented immune overreaction, called regulatory T cells, or Tregs. Without these cells, the mice were unusually prone to inflammatory disease.

Honda wanted to know which of the many intestinal species might induce these suppressor cells. Soon after Sokol identified the anti-inflammatory effects of *F. prausnitzii*, Honda began whittling away at the gut microbiota of mice by treating them with narrow-spectrum antibiotics. The animals’ Tregs declined after a course of vancomycin. With their ability to restrain their immune reaction hobbled, the mice became highly susceptible to colitis, the rodent version of inflammatory bowel disease and allergic diarrhea. Honda found he could restore the Tregs and immune equilibrium of the mice just by reinstating 46 native clostridial strains.

Honda repeated the exercise with human-adapted microbes obtained from a healthy lab member. He extracted just 17 clostridial species this time that, in mice, could induce a full repertoire of Tregs and prevent inflammation. These human-adapted microbes specialized in nudging the immune system away from inflammatory disease. They came from branches of the clostridial group labeled clusters IV, XIVa and XVIII. *F. prausnitzii* belongs to cluster IV.

Vedanta Biosciences recently formed to try to turn Honda’s 17-strain “clostridial cocktail” into a treatment for inflammatory disease. If the company’s efforts are successful, it could signal the arrival of the next generation of probiotics—human-adapted microbes to treat immune-mediated disease—and all derived from one member of Honda’s lab. As always, it is unclear if what works in lab mice will translate to humans. Sokol has his doubts. He recently identified a type of regulatory T cell that is unique to humans and that is deficient in people with inflammatory bowel disease. He questions if Honda’s cocktail, which has been developed in mice, will activate these cells in people.

TROUBLE WITH ANTIBIOTICS

EVEN IF THE COCKTAIL falls short, Honda’s meticulous demonstration of a link between antibiotics and vulnerability to inflammatory disease has raised a troubling question. A number of studies have found a small but significant correlation between the early-life use of antibiotics and the later development of inflammatory disorders, including asthma, inflammatory bowel disease and, more recently, colorectal cancer and childhood obesity. One explanation for this association might be that sickly people take more antibiotics. Antibiotics are not the cause, in other words, but the result of preexisting ill health.

Honda’s studies suggest another explanation: antibiotics may deplete the very bacteria that favorably calibrate the immune system, leaving it prone to overreaction. Brett Finlay, a microbiologist at the University of British Columbia, has explored this possibility explicitly. Early-life vancomycin treatment of mice increased the animals’ risk of asthma later, he found, in part by depleting those very same clostridial bacteria identified by Honda. The corresponding population of suppressor cells collapsed. And the animals became less able to restrain their immune responses when encountering allergens later.

These dynamics may also apply to other diseases. Earlier this year Cathryn Nagler, an immunologist at the University of Chicago,

The Gene-Microbe Link

Evidence that genes shape the microbiome may point to new treatments for common diseases *By Ruth E. Ley*

The ecology of the gut microbiome may trigger or contribute to a variety of diseases, including autoimmune disorders and obesity, research suggests. Factors such as early environment, diet and antibiotic exposure have a lot to do with why people differ from one another in the composition of their microbiomes. But specific gene variants are also linked to greater risks of developing many of these diseases. Do your genes act on your microbiome, which in turn promotes disease?

One way researchers have addressed this question is to pick specific genes that are good candidates—for instance, those with a strong link to a disease that also has a microbiome link—and examine whether people who carry mutations that are known to increase the risk of a certain disease also have microbiomes that differ from those who do not have the mutations. A team led by Dan Frank at the University of Colorado Denver took this approach and revealed that specific variants of the *NOD2* gene that confer a high risk of developing inflammatory bowel disease to their carriers are also associated with an altered intestinal microbiome.

A powerful and broader way to look for an effect of human genetic variation on the microbiome is to compare twins. Identical twins share nearly 100 percent of their genes; fraternal twins, 50 percent. Co-twins are raised together, so the environmental effects on their microbiomes should be about the same. If the microbiomes of the identical twins are more alike within

a twinship than those of the fraternal twins, we can conclude that genes have played a role. If variation within twinships of each kind is about the same, we can say a shared genome has had no additional effect.

Early twins studies were based on fewer than 50 twin pairs and could not detect any greater similarities in the microbiomes of identical twins compared with those of fraternal twins. But recent work my laboratory at Cornell University conducted with researchers at King's College London compared nearly 500 twin pairs, a sample size sufficient to show a marked genetic effect on the relative abundance of a specific set of gut microbes. Furthermore, so-called heritable microbes—the bacteria most influenced by host genetics—were more abundant in lean twins than obese ones.

Experiments in germ-free mice showed that one gut bacterium in particular, *Christensenella minuta*, can influence the phenotype—the composite of observable characteristics or traits—of the host. Germ-free mice live in sterile bubbles—and they are very skinny. When they are given a microbiome in the form of a fecal transplant from a human donor, however, they plump up within a day or two because the bacteria help them digest their food and develop a proper metabolism. We found that if *C. minuta* was added to the feces of an obese human donor, the recipient mice were thinner than when *C. minuta* was not added. Results showing *C. minuta* has an effect of controlling fat gain in the mouse match data that reveal lean people have a greater abundance of *C. minuta* in their gut than obese people.

This is evidence that a person's genes can influence the gut microbiome's composition and in turn can shape the individual's phenotype. Further work will show what specific genes are involved as well as how the microbiome may be reshaped to reduce risk of developing chronic inflammatory diseases within the context of a person's genotype, suggesting potential new approaches to treating obesity-related diseases.

Ruth E. Ley is an associate professor of molecular biology and genetics at Cornell University.



A powerful way to look for an effect of human genetic variation on the microbiome is to compare twins.

knocked out the clostridial bacteria with antibiotics and then fed the animals peanut protein. Without those microbes and their corresponding Tregs present, the protein leaked through the gut barrier into circulation, prompting the rodent version of a food allergy. She could prevent the sensitization just by introducing those clostridial bacteria.

One key difference between mice with and without the clostridial clusters was how many mucus-secreting cells they possessed. Animals that harbored the clostridial clusters had more. That may have far-reaching consequences. Mucus, scientists are finding, contains compounds that repel certain microbes, maintaining a tiny distance between them and us. But it also carries food for other bacteria—complex, fermentable sugars that resemble those found in breast milk. Lora Hooper, a microbiologist at the University of Texas Southwestern Medical Center in Dallas, calls this dual function the “carrot” and the “stick.” Mucus serves both as an antimicrobial repellent and a growth medium for friendly bacteria.

This phenomenon matters for several reasons. As Nagler’s experiments suggest, one way these clostridial clusters may promote gut health and a balanced immune system is by ensuring a healthy flow of mucus. Just as those elephants help to maintain the African savanna, these microbes may favorably shape the greater gut ecosystem by stimulating secretion of the sugars other friendly microbes graze on.

Conversely, scientists observe defects in the mucous layer in other disorders, particularly inflammatory bowel disease, where these clostridial bacteria are often depleted. The question has always been which comes first: defects in mucus secretion and the selection of an aberrant community of microbes or acquisition of an aberrant community of microbes that thins the mucous layer and increases vulnerability to disease? Both factors may work together.

In 2011 scientists at the University of Colorado Boulder sampled people with variants of a gene called *NOD2* associated with inflammatory bowel disease. No one quite understands how these variants of the gene, which codes for a microbial sensor, increase the risk of disease. Study participants included people both with and without disease. Those suffering from inflammatory bowel disease had reduced counts of clostridial bacteria, the scientists found. But more surprising, people who did not have disease but who carried the predisposing *NOD2* variants also had a relative depletion of clostridial clusters. Their microbial communities seemed positioned closer to a diseaselike state.

The study seems to highlight the role of genes in determining the composition of gut microbiota and the vulnerability to Crohn’s. But epidemiological surveys complicate the picture. A number of studies over the years have linked having fewer sanitary amenities in childhood with a lower risk of inflammatory bowel disease in adulthood. And a 2014 study from Aarhus University in Denmark found that among northern Europeans, growing up on a farm with livestock—another microbially enriched environment—halved the risk of being stricken with inflammatory bowel disease in adulthood.

These patterns suggest that perhaps by seeding the gut microbiota early in life or by direct modification of the immune system the environment can affect our risk of inflammatory bowel disease despite the genes we carry. And they raise the question of what proactive steps

those of us who do not live on farms can take to increase our chances of harboring a healthy mix of microbes.

THE IMPORTANCE OF FIBER

ONE OF THE MORE SURPRISING discoveries in recent years is how much the gut microbiota of people living in North America differs from those of people living in rural conditions in Africa and South America. The microbial mix in North America is geared to digesting protein, simple sugars and fats, whereas the mix in rural African and Amazonian environments is far more diverse and geared to fermenting plant fiber. Some think that our hunter-gatherer ancestors harbored even greater microbial diversity in their guts. If we accept the gut microbiota of people in rural Africa and South America as proxies for those that prevailed before the industrial revolution, then, says Justin L. Sonnenburg, a microbiologist at Stanford University, the observed differences suggest North Americans and other Westernized populations have veered into evolutionarily novel territory.

What troubles Sonnenburg about this shift is that the bacteria that seem most anti-inflammatory—including the clostridial clusters—often specialize in fermenting soluble fiber. Fermentation produces various metabolites, including butyrate, acetate and propionate—some of the substances that produce underarm odor. Various rodent studies suggest that these metabolites, called short-chain fatty acids, can induce Tregs and calibrate immune function in ways that, over a lifetime, may prevent inflammatory disease. Fermentation by-products may be one way our gut microbes communicate with our bodies. One takeaway is to “feed your Tregs more fiber,” as University of Oxford immunologist Fiona Powrie put it last year in the journal *Science*.

Yet the seeming importance of these metabolites has others puzzled. Many bacteria produce these short-chain fatty acids, and yet only a few microbes seem potentially anti-inflammatory. So although production of these metabolites may be a prerequisite for microbes that favorably tweak the immune system, says Sarkis Mazmanian, a microbiologist at the California Institute of Technology, it is insufficient to explain why some bacteria are more anti-inflammatory than others. Other characteristics, such as how close they live to the gut lining or the molecules they use to prod the host immune system, must also play a role, he says.

There is, however, an issue of sheer quantity. Some hunter-gatherers consumed up to 10 times as much soluble fiber as modern populations, and their bodies likely were flooded with far more fermentation by-products. Our fiber-poor modern diet may have weakened that signal, producing a state of “simmering hyperactivity,” Sonnenburg says, and predisposing us to the “plagues” of civilization. He calls this problem “starving our microbial self.” We may not be adequately feeding some of the most important members of our microbiota.

Mouse experiments support the idea. Diets high in certain fats and sugars deplete anti-inflammatory bacteria, thin the mucous layer and foster systemic inflammation. Potentially dangerous opportunists bloom. In one intervention on human volunteers, University of California, San Francisco, microbiologist Peter Turnbaugh found that switching to a high-fat, high-protein diet spurred an expansion of bile-tolerant bacteria, one of which, *Bilophila wadsworthia*, has been

Your Microbes at Work: Fiber Fermenters Keep Us Healthy

The gut houses trillions of microbes. They eat what you eat. Many specialize in fermenting the soluble fiber in legumes, grains, fruits and vegetables. Certain microbial species are adept at colonizing the mucous layer of the gut. Mucus contains antimicrobial substances that keep the microbiota at a slight distance. But it also contains sugars such as those found in breast milk. Some microbes, often the same ones that specialize in fermenting fiber, can use these sugars as sustenance when other food is not available. The by-products of fiber fermentation nourish cells lining the colon. Some by-products pass into the circulation and may calibrate our immune system in a way that prevents inflammatory disorders such as asthma and Crohn's disease.

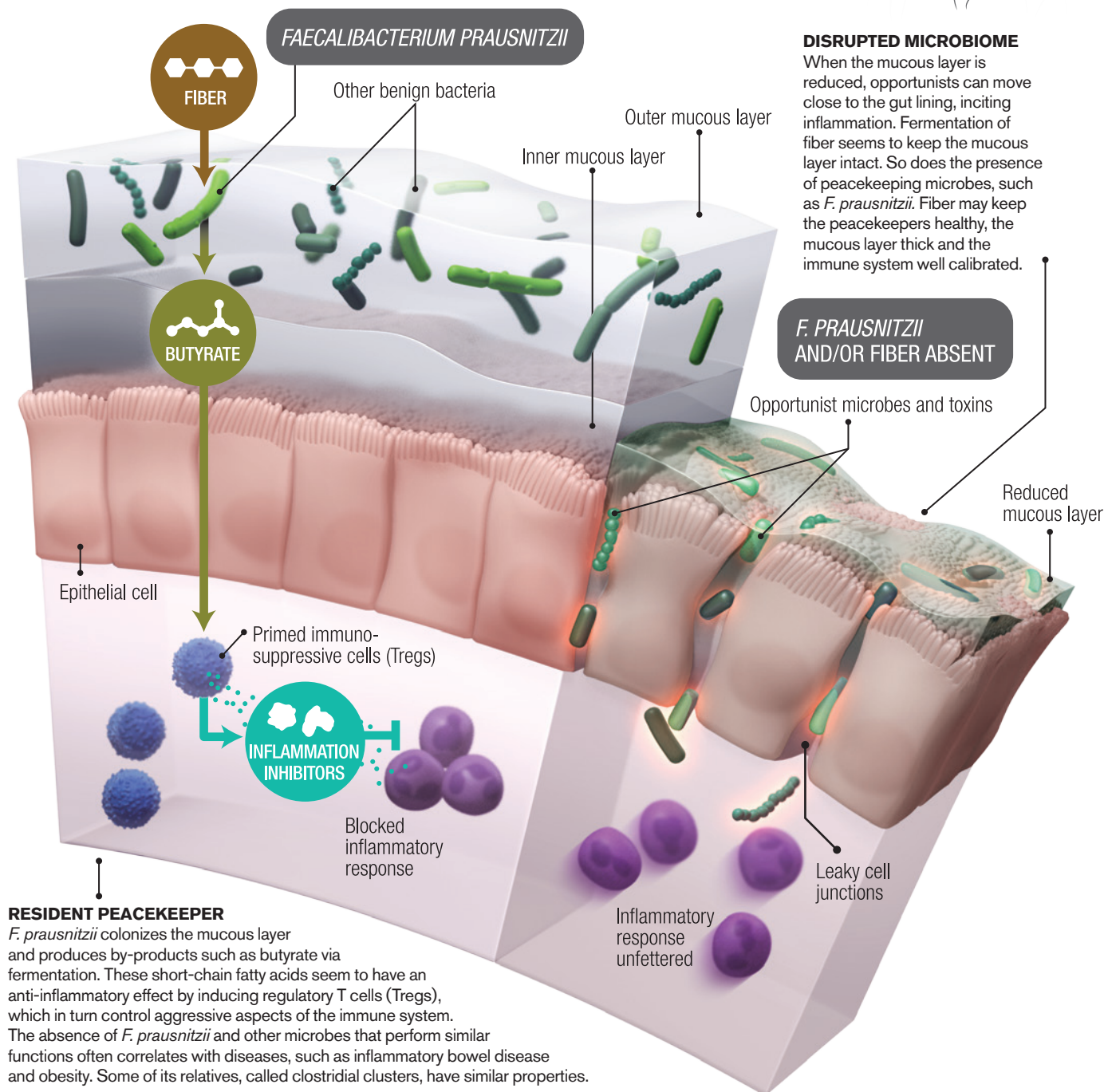
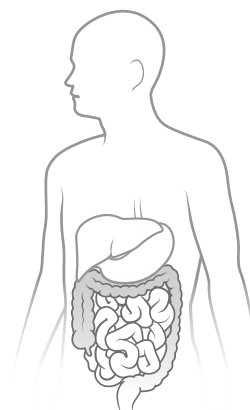


ILLUSTRATION BY AXS BIOMEDICAL ANIMATION STUDIO.
 SOURCE: "FEED YOUR TREGS MORE FIBER," BY JULIA BOLLRATH
 AND FIONA POWRIE, IN *SCIENCE*, VOL. 341, AUGUST 2, 2013

linked to inflammatory bowel disease. On the other hand, preventing this skewing of the microbial self does not seem that difficult. In rodents, adding fermentable fiber to a diet otherwise high in fat keeps the “good” microbes happy, the mucous layer healthy and the gut barrier intact, and it prevents systemic inflammation. Taken together, these studies suggest that it is not only what is in your food that matters for your health but also what is missing.

The human studies are even more intriguing. Mounting evidence suggests that the systemic inflammation observed in obesity does not just result from the accumulation of fat but contributes to it. Scientists at Catholic University of Louvain in Belgium recently showed that adding inulin, a fermentable fiber, to the diet of obese women increased counts of *F. prausnitzii* and other clostridial bacteria and reduced that dangerous systemic inflammation. Weight loss was minor, but later analysis of this and two similar studies revealed that the intervention worked best on patients who, at the outset, already harbored clostridial clusters IV, IX and XIVa—some of the same clusters represented in Honda’s cocktail. Those without the bacteria did not benefit, which suggests that once species disappear from the “microbial organ,” the associated functions might also vanish. These individuals might not require ecosystem engineering so much as an ecosystem restoration.

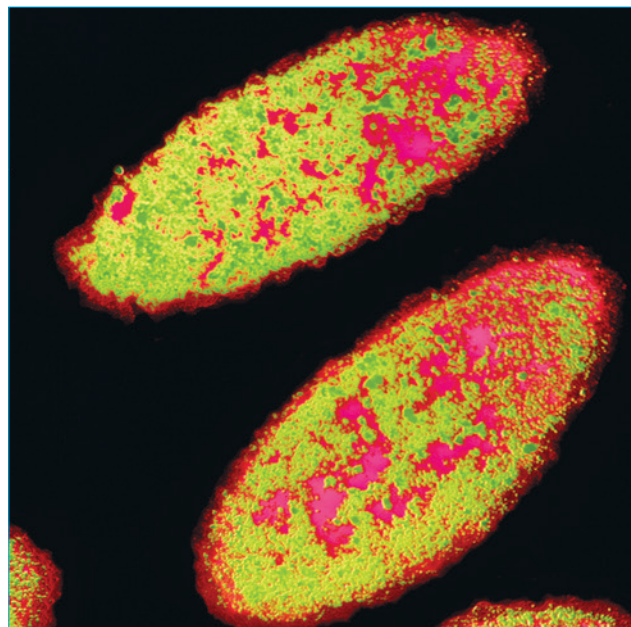
That possibility has also been tested. Several years ago Max Nieuwdorp, a gastroenterologist at the Academic Medical Center in Amsterdam, transplanted microbes from lean donors to patients recently diagnosed with metabolic syndrome, a cluster of symptoms that often predicts type 2 diabetes. The recipients saw improvements in insulin sensitivity and an enrichment of their microbiota, including among those clostridial species. But six months after the transplant the patients had relapsed, metabolic improvements had faded and their microbes had reverted to their original states.

To Sonnenburg, this outcome suggests that the dance between human host and microbial community has considerable momentum. Removing the “diseased” ecosystem and installing a new one may not overcome the inertia. The gut immune system may simply mold the new community in the image of the old. That may explain why fecal transplants, which effectively vanquish *C. difficile*—associated diarrhea, have so far failed to treat inflammatory bowel disease. The former is caused by a single opportunist; the latter may be driven by an out-of-whack ecosystem and our response to the microbial derangement.

To overcome the inertia, Sonnenburg foresees treating the host and the microbiota simultaneously. The idea has not been tested, but he imagines clearing out the microbiota, perhaps with antibiotics, followed by immunosuppressants to quiet the patient’s immune system and allow healing. Only then might the new community of microbes stick and successfully recalibrate the immune system.

EVOLUTION OF MOBILITY

WHEN ANIMAL LIFE EXPLODED some 800 million years ago, microbes had already existed on Earth for maybe three billion years. A major innovation in animal evolution was the gut—a tube that takes nutrients in one end and expels waste from the other. It is even possible, argues Margaret McFall-Ngai, a microbiologist at the University of



DANGEROUS OPPORTUNIST: *Bilophila wadsworthia*, a species of bacterium linked to inflammatory bowel disease, bloomed in the microbiota of human volunteers fed a high-fat, high-protein diet in a recent experiment.

Wisconsin–Madison, that microbes drove the evolution of the gut directly. Plants only succeeded in colonizing land when they had developed relationships with microbes that helped them extract vital nutrients from soil. Perhaps one evolutionary innovation of animals was to scoop up the microbial communities necessary for survival and to take them along for the ride, achieving mobility.

Mucus may be one way the human gut selects for these microbes. Only co-adapted bacteria, Sonnenburg thinks, can metabolize the complex sugars it contains. A cornerstone of this symbiosis may be the simple imperative of acquiring nutrients in a world of scarcity. We hunt and gather the goods; the microbes ferment what we cannot digest, taking a cut in the process and keeping pathogens at bay. Our immune systems quiet down when they receive signals, conveyed partly in microbial metabolites, indicating that the right microbes are in place.

The field of gut microbiome research has already moved from the idea of describing the core species to identifying the core ecological functions various microbes perform. Many potential species may fulfill any given role. Now another concept may be emerging, which might be called the keystone relationship. “The interaction between fiber and microbes that consume it,” Sonnenburg says, “is the fundamental keystone interaction that everything else is built on in the gut.” It may lie at the heart of the symbiotic pact between microbes and humans.

Moises Velasquez-Manoff is author of *An Epidemic of Absence: A New Way of Understanding Allergies and Autoimmune Diseases*. His work has appeared in the *New York Times*, *Mother Jones* and *Nautilus*.



THE NOTION THAT THE STATE of our gut governs our state of mind dates back more than 100 years. Many 19th- and early 20th-century scientists believed that accumulating wastes in the colon triggered a state of “auto-intoxication,” whereby poisons emanating from the gut produced infections that were in turn linked with depression, anxiety and psychosis. Patients were treated with colonic purges and even bowel surgeries until these practices were dismissed as quackery.

The ongoing exploration of the human microbiome promises to bring the link between the gut and the brain into clearer focus. Scientists are increasingly convinced that the vast assemblage of microfauna in our intestines may have a major impact on our state of mind. The gut-brain axis seems to be bidirectional—the brain acts on gastrointestinal and immune functions that help to shape the gut’s microbial makeup, and gut microbes make neuroactive compounds, including neurotransmitters and metabolites that also act on the brain. These interactions could occur in various ways: microbial compounds communicate via the vagus nerve, which connects the brain and the digestive tract, and microbially derived metabolites interact with the immune system, which maintains its own communication with the brain. Sven Pettersson, a microbiologist at the Karolinska Institute in Stockholm, has recently shown that gut microbes help to control leakage through both the intestinal lining and the blood-brain barrier, which ordinarily protects the brain from potentially harmful agents.

Microbes may have their own evolutionary reasons for communicating with the brain. They need us to be social, says John Cryan, a neuroscientist at University College Cork in Ireland, so that they can spread through the human population. Cryan’s research shows that when bred in sterile conditions, germ-free mice lacking in intestinal microbes also lack an ability to recognize other mice with whom they interact. In other studies, disruptions of the microbiome induced mice behavior that mimics human anxiety, depression and even autism. In some cases, scientists restored more normal behavior by treating their test subjects with certain strains of benign bacteria. Nearly all the data so far are limited to mice, but Cryan believes the findings provide fertile ground for developing analogous compounds, which he calls psychobiotics, for humans. “That dietary treatments could be used as either adjunct or sole therapy

for mood disorders is not beyond the realm of possibility,” he says.

PERSONALITY SHIFTS

SCIENTISTS USE germ-free mice to study how the lack of a microbiome—or selective dosing with particular bacteria—alters behavior and brain function, “which is something we could never do in people,” Cryan says. Entire colonies of germ-free mice are bred and kept in isolation chambers, and the technicians who handle them wear full bodysuits, as if they were in a biohazard facility. As with all mice research, extrapolating results to humans is a big step. That is especially true with germ-free mice because their brains and immune systems are underdeveloped, and they tend to be more hyperactive and daring than normal mice.

A decade ago a research team led by Nobuyuki Sudo, now a professor of internal medicine at Kyushu University in Japan, restrained germ-free mice in a narrow tube for up to an hour and then measured their stress hormone output. The amounts detected in the germ-free animals were far higher than those measured in normal control mice exposed to the same restraint. These hormones are released

by the hypothalamic-pituitary-adrenal axis, which in the germ-free mice was clearly dysfunctional. But more important, the scientists also found they could induce more normal hormonal responses simply by pretreating the animals with a single microbe: a bacterium called *Bifidobacterium infantis*. This finding showed for the first time that intestinal microbes could influence stress responses in the brain and hinted at the possibility of using probiotic treatments to affect brain function in beneficial ways. “It really got the field off the ground,” says Emeran Mayer, a gastroenterologist and director of the Center for Neurobiology of Stress at the University of California, Los Angeles.

Meanwhile a research team at McMaster University in Ontario led by microbiologist Premysl Bercik and gastroenterologist Stephen Collins discovered that if they colonized the intestines of one strain of

MENTAL HEALTH

Thinking from the Gut

The microbiome may yield a new class of psychobiotics for the treatment of anxiety, depression and other mood disorders

By Charles Schmidt

The Diverse Microbiome of the Hunter-Gatherer



The Hadza of Tanzania offer a snapshot of the co-adaptive capacity of the gut ecosystem

By Stephanie L. Schnorr

We tend to forget that modern humanity is largely sheltered from the last vestiges of wild untamed Earth and that our way of life bears little resemblance to how our ancestors lived during 90 percent of human history. We have lost nearly all trace of these former selves—and, worse, have marginalized the few remaining humans who retain their hunter-gatherer identity. In Tanzania, tribes of

wandering foragers called the Hadza, who have lived for thousands of years in the East African Rift Valley ecosystem, tell us an immense and precious story about how humans, together with their microbial evolutionary partners, are adapted to live and thrive in a complex natural environment.

Ongoing research with the Hadza to characterize the hunter-gatherer-microbiome relationship has yielded not only insight into the co-adaptive capacity of this microbial ecosystem but also a profound

appreciation for how versatile human life can be. The microbiome is central to our biology. It mediates the interaction and exchange of information across host-environment thresholds such as the mouth, skin and gut.

The strength and importance of this mediation are borne out in the Hadza gut microbiota. Their microbiome harbors incredibly high taxonomic diversity, indicating great ecosystem stability and flexibility. It is capable of withstanding the perpetual presence of parasites and pathogens and can respond to fluctuations in diet caused by an unpredictable and seasonally dependent food supply. Interestingly, bacterial taxonomic abundance is different in Hadza men and women. Because of the sexual division of labor in Hadza society, men and women tend to consume more of their respective foraged food resources. The women primarily collect and eat tubers and other plant foods. As a result, it appears that women carry more bacteria to help process the plant fiber in their diets. This difference has direct implications for how the gut microbiota may enable Hadza

germ-free mice with bacteria taken from the intestines of another mouse strain, the recipient animals would take on aspects of the donor's personality. Naturally timid mice would become more exploratory, whereas more daring mice would become apprehensive and shy. These tendencies suggested that microbial interactions with the brain could induce anxiety and mood disorders.

Bercik and Collins segued into gut-brain research from their initial focus on how the microbiome influences intestinal illnesses. People who suffer from these conditions often have co-occurring psychiatric problems such as anxiety and depression that cannot be fully explained as an emotional reaction to being sick. By colonizing germ-free mice with the bowel contents of people with irritable bowel syndrome, which induces constipation, diarrhea, pain and low-grade inflammation but has no known cause, the McMaster's team reproduced many of the same gastrointestinal symptoms. The animals developed leaky intestines, their immune systems activated, and they produced a barrage of pro-inflammatory metabolites, many with known nervous system effects. Moreover, the mice also displayed anxious behavior, as indicated in a test of their willingness to step down from a short raised platform.

AUTISM CONNECTION?

SCIENTISTS HAVE ALSO BEGUN to explore the microbiome's potential role in autism. In 2007 the late Paul Patterson, a neuroscientist and developmental biologist at the California Institute of Technology, was

intrigued by epidemiological data showing that women who suffer from a high, prolonged fever during pregnancy are up to seven times more likely to have a child with autism. These data suggested an alternative cause for autism besides genetics. To investigate, Patterson induced flu-like symptoms in pregnant mice with a viral mimic: an immunostimulant called polyinosinic:polycytidylic acid, or poly(I:C). He called this the maternal immune activation (MIA) model.

The offspring of Patterson's MIA mice displayed all three of the core features of human autism: limited social interactions, a tendency toward repetitive behavior and reduced communication, which he assessed by using a special microphone to measure the length and duration of their ultrasonic vocalizations. In addition, the mice had leaky intestines, which was important because anywhere from 40 to 90 percent of all children with autism suffer from gastrointestinal symptoms.

Then Caltech microbiologist Sarkis Mazmanian and his doctoral student Elaine Hsiao discovered that MIA mice also have abnormal microbiomes. Specifically, two bacterial classes—Clostridia and Bacteroidia—were far more abundant in the MIA offspring than in normal mice. Mazmanian acknowledges that these imbalances may not be the same as those in humans with autism. But the finding was compelling, he says, because it suggested that the behavioral state of the MIA mice—and perhaps by extension autistic behavior in humans—might be rooted in the gut rather than the brain. “That raised a provocative question,” Mazmanian says. “If we treated gastrointestinal

women to obtain adequate nutrition for fertility and reproductive success, despite a resource-limited environment. Through our work with the Hadza, we have been able to contribute to mounting evidence that human microbiota exerts a powerful influence on host health and survival, especially in natural fertility- and subsistence-based populations.

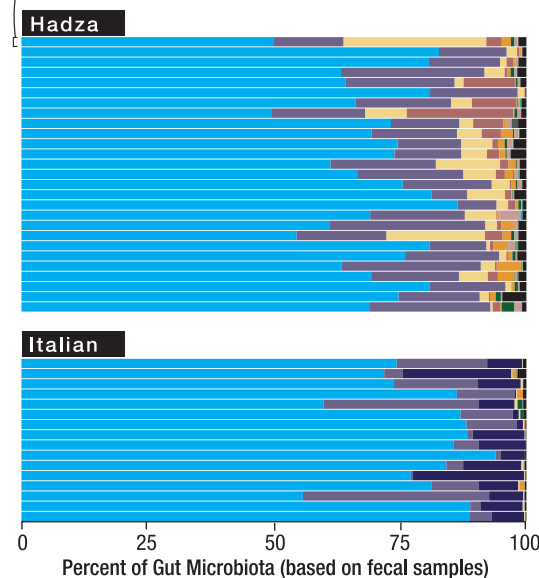
Comparative analysis of the gut microbiota of hunter-gatherers with that of Westernized industrial populations is also beginning to yield important insights. The microbial diversity in industrial groups is far below that of the Hadza, as well as those of other rural farming communities in Burkina Faso, Malawi and South Africa. Whereas a reduction in diversity may not seem ideal, it is the predictable response of an ecosystem facing a narrow range of selective pressures and is therefore no less adaptive. Some technological interventions, such as hypersanitation, consumption of refined foods and habitual use of antibiotics, have had a dramatic impact over time on the functional role of the microbiome in industrial populations. These

aspects of a Westernized way of life have to a large extent displaced much of the original mutualistic functions of the microbiome in stabilizing our bodies against foreign microorganisms, allowing us to digest unprocessed foods and helping train our immune system to effectively fight disease.

We are just beginning to understand how the microbiome evolves over our lifetimes as a dynamic and mutualistic ecosystem that helps to facilitate human health. Thanks to the Hadza, we know that ancient human hunter-gatherers must have maintained a direct and persistent interface with the natural environment. As a result, the ancestral human microbiome was almost certainly a taxonomically diverse community, providing the functional flexibility that accompanied global colonization and is our adaptive legacy.

Stephanie L. Schnorr is a Ph.D. candidate in the Research Group on Plant Foods in Hominin Dietary Ecology at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany.

Each horizontal bar represents an individual in the study group
Each color represents a different microbe phylum (■ = Firmicutes)



TAXONOMIC TREASURE TROVE:

A survey of fecal microbiota of 43 subjects revealed a more varied mix of gut bacteria phyla among Hadza hunter-gatherers compared with urban Italians.

symptoms in the mice, would we see changes in their behavior?"

Mazmanian and Hsiao investigated by dosing the animals with a microbe known for its anti-inflammatory properties, *Bacteroides fragilis*, which also protects mice from experimentally induced colitis. Results showed that the treatment fixed intestinal leaks and restored a more normal microbiota. It also mitigated the tendency toward repetitive behavior and reduced communication. Mazmanian subsequently found that *B. fragilis* reverses MIA deficits even in adult mice. "So, at least in this mouse model, it suggests features of autism aren't hardwired—they're reversible—and that's a huge advance," he says.

LIMITS OF RESEARCH

THE HUMAN GUT MICROBIOME evolved to help us in myriad ways: Gut microbes make vitamins, break dietary fiber into digestible short-chain fatty acids and govern normal functions in the immune system. Probiotic treatments such as yogurt supplemented with beneficial strains of bacteria are already being used to help treat some gastrointestinal disorders, such as antibiotic-induced diarrhea. But there are little data about probiotic effects on the human brain.

In a proof-of-concept study Mayer and his colleagues at U.C.L.A. uncovered the first evidence that probiotics ingested in food can alter human brain function. The researchers gave healthy women yogurt twice a day for a month. Then brain scans using functional magnetic resonance imaging were taken as the women were shown pictures of

actors with frightened or angry facial expressions. Normally, such images trigger increased activity in emotion-processing areas of the brain that leap into action when someone is in a state of heightened alert. Anxious people may be uniquely sensitive to these visceral reactions. But the women on the yogurt diet exhibited a less "reflexive" response, "which shows that bacteria in our intestines really do affect how we interpret the world," says gastroenterologist Kirsten Tillisch, the study's principal investigator. Mayer cautions that the results are rudimentary. "We simply don't know yet if probiotics will help with human anxiety," he says. "But our research is moving in that direction."

Strains of *Bifidobacterium*, which is common in the gut flora of many mammals, including humans, have generated the best results so far. Cryan recently published a study in which two varieties of *Bifidobacterium* produced by his lab were more effective than escitalopram (Lexapro) at treating anxious and depressive behavior in a lab mouse strain known for pathological anxiety. Although Cryan is optimistic that such findings may point the way to the development of psychobiotics, he is wary of hype. "We still need a lot more research into the mechanisms by which gut bacteria interact with the brain," he says.

Charles Schmidt is a recipient of the National Association of Science Writers' Science in Society Journalism Award. His work has appeared in *Science*, *Nature Biotechnology*, *Nature Medicine* and *the Washington Post*.

The Diverse Microbiome of the Hunter-Gatherer



The Hadza of Tanzania offer a snapshot of the co-adaptive capacity of the gut ecosystem

By Stephanie L. Schnorr

We tend to forget that modern humanity is largely sheltered from the last vestiges of wild untamed Earth and that our way of life bears little resemblance to how our ancestors lived during 90 percent of human history. We have lost nearly all trace of these former selves—and, worse, have marginalized the few remaining humans who retain their hunter-gatherer identity. In Tanzania, tribes of

wandering foragers called the Hadza, who have lived for thousands of years in the East African Rift Valley ecosystem, tell us an immense and precious story about how humans, together with their microbial evolutionary partners, are adapted to live and thrive in a complex natural environment.

Ongoing research with the Hadza to characterize the hunter-gatherer-microbiome relationship has yielded not only insight into the co-adaptive capacity of this microbial ecosystem but also a profound

appreciation for how versatile human life can be. The microbiome is central to our biology. It mediates the interaction and exchange of information across host-environment thresholds such as the mouth, skin and gut.

The strength and importance of this mediation are borne out in the Hadza gut microbiota. Their microbiome harbors incredibly high taxonomic diversity, indicating great ecosystem stability and flexibility. It is capable of withstanding the perpetual presence of parasites and pathogens and can respond to fluctuations in diet caused by an unpredictable and seasonally dependent food supply. Interestingly, bacterial taxonomic abundance is different in Hadza men and women. Because of the sexual division of labor in Hadza society, men and women tend to consume more of their respective foraged food resources. The women primarily collect and eat tubers and other plant foods. As a result, it appears that women carry more bacteria to help process the plant fiber in their diets. This difference has direct implications for how the gut microbiota may enable Hadza

germ-free mice with bacteria taken from the intestines of another mouse strain, the recipient animals would take on aspects of the donor's personality. Naturally timid mice would become more exploratory, whereas more daring mice would become apprehensive and shy. These tendencies suggested that microbial interactions with the brain could induce anxiety and mood disorders.

Bercik and Collins segued into gut-brain research from their initial focus on how the microbiome influences intestinal illnesses. People who suffer from these conditions often have co-occurring psychiatric problems such as anxiety and depression that cannot be fully explained as an emotional reaction to being sick. By colonizing germ-free mice with the bowel contents of people with irritable bowel syndrome, which induces constipation, diarrhea, pain and low-grade inflammation but has no known cause, the McMaster's team reproduced many of the same gastrointestinal symptoms. The animals developed leaky intestines, their immune systems activated, and they produced a barrage of pro-inflammatory metabolites, many with known nervous system effects. Moreover, the mice also displayed anxious behavior, as indicated in a test of their willingness to step down from a short raised platform.

AUTISM CONNECTION?

SCIENTISTS HAVE ALSO BEGUN to explore the microbiome's potential role in autism. In 2007 the late Paul Patterson, a neuroscientist and developmental biologist at the California Institute of Technology, was

intrigued by epidemiological data showing that women who suffer from a high, prolonged fever during pregnancy are up to seven times more likely to have a child with autism. These data suggested an alternative cause for autism besides genetics. To investigate, Patterson induced flu-like symptoms in pregnant mice with a viral mimic: an immunostimulant called polyinosinic:polycytidylic acid, or poly(I:C). He called this the maternal immune activation (MIA) model.

The offspring of Patterson's MIA mice displayed all three of the core features of human autism: limited social interactions, a tendency toward repetitive behavior and reduced communication, which he assessed by using a special microphone to measure the length and duration of their ultrasonic vocalizations. In addition, the mice had leaky intestines, which was important because anywhere from 40 to 90 percent of all children with autism suffer from gastrointestinal symptoms.

Then Caltech microbiologist Sarkis Mazmanian and his doctoral student Elaine Hsiao discovered that MIA mice also have abnormal microbiomes. Specifically, two bacterial classes—Clostridia and Bacteroidia—were far more abundant in the MIA offspring than in normal mice. Mazmanian acknowledges that these imbalances may not be the same as those in humans with autism. But the finding was compelling, he says, because it suggested that the behavioral state of the MIA mice—and perhaps by extension autistic behavior in humans—might be rooted in the gut rather than the brain. “That raised a provocative question,” Mazmanian says. “If we treated gastrointestinal

women to obtain adequate nutrition for fertility and reproductive success, despite a resource-limited environment. Through our work with the Hadza, we have been able to contribute to mounting evidence that human microbiota exerts a powerful influence on host health and survival, especially in natural fertility- and subsistence-based populations.

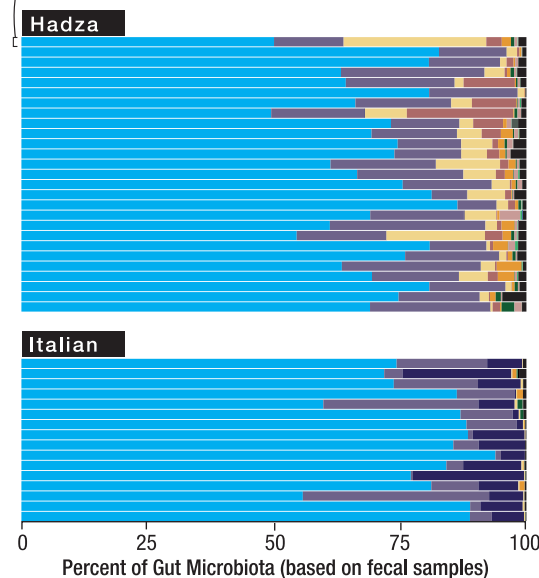
Comparative analysis of the gut microbiota of hunter-gatherers with that of Westernized industrial populations is also beginning to yield important insights. The microbial diversity in industrial groups is far below that of the Hadza, as well as those of other rural farming communities in Burkina Faso, Malawi and South Africa. Whereas a reduction in diversity may not seem ideal, it is the predictable response of an ecosystem facing a narrow range of selective pressures and is therefore no less adaptive. Some technological interventions, such as hypersanitation, consumption of refined foods and habitual use of antibiotics, have had a dramatic impact over time on the functional role of the microbiome in industrial populations. These

aspects of a Westernized way of life have to a large extent displaced much of the original mutualistic functions of the microbiome in stabilizing our bodies against foreign microorganisms, allowing us to digest unprocessed foods and helping train our immune system to effectively fight disease.

We are just beginning to understand how the microbiome evolves over our lifetimes as a dynamic and mutualistic ecosystem that helps to facilitate human health. Thanks to the Hadza, we know that ancient human hunter-gatherers must have maintained a direct and persistent interface with the natural environment. As a result, the ancestral human microbiome was almost certainly a taxonomically diverse community, providing the functional flexibility that accompanied global colonization and is our adaptive legacy.

Stephanie L. Schnorr is a Ph.D. candidate in the Research Group on Plant Foods in Hominin Dietary Ecology at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany.

Each horizontal bar represents an individual in the study group
Each color represents a different microbe phylum (■ = Firmicutes)



TAXONOMIC TREASURE TROVE:

A survey of fecal microbiota of 43 subjects revealed a more varied mix of gut bacteria phyla among Hadza hunter-gatherers compared with urban Italians.

symptoms in the mice, would we see changes in their behavior?"

Mazmanian and Hsiao investigated by dosing the animals with a microbe known for its anti-inflammatory properties, *Bacteroides fragilis*, which also protects mice from experimentally induced colitis. Results showed that the treatment fixed intestinal leaks and restored a more normal microbiota. It also mitigated the tendency toward repetitive behavior and reduced communication. Mazmanian subsequently found that *B. fragilis* reverses MIA deficits even in adult mice. "So, at least in this mouse model, it suggests features of autism aren't hardwired—they're reversible—and that's a huge advance," he says.

LIMITS OF RESEARCH

THE HUMAN GUT MICROBIOME evolved to help us in myriad ways: Gut microbes make vitamins, break dietary fiber into digestible short-chain fatty acids and govern normal functions in the immune system. Probiotic treatments such as yogurt supplemented with beneficial strains of bacteria are already being used to help treat some gastrointestinal disorders, such as antibiotic-induced diarrhea. But there are little data about probiotic effects on the human brain.

In a proof-of-concept study Mayer and his colleagues at U.C.L.A. uncovered the first evidence that probiotics ingested in food can alter human brain function. The researchers gave healthy women yogurt twice a day for a month. Then brain scans using functional magnetic resonance imaging were taken as the women were shown pictures of

actors with frightened or angry facial expressions. Normally, such images trigger increased activity in emotion-processing areas of the brain that leap into action when someone is in a state of heightened alert. Anxious people may be uniquely sensitive to these visceral reactions. But the women on the yogurt diet exhibited a less "reflexive" response, "which shows that bacteria in our intestines really do affect how we interpret the world," says gastroenterologist Kirsten Tillisch, the study's principal investigator. Mayer cautions that the results are rudimentary. "We simply don't know yet if probiotics will help with human anxiety," he says. "But our research is moving in that direction."

Strains of *Bifidobacterium*, which is common in the gut flora of many mammals, including humans, have generated the best results so far. Cryan recently published a study in which two varieties of *Bifidobacterium* produced by his lab were more effective than escitalopram (Lexapro) at treating anxious and depressive behavior in a lab mouse strain known for pathological anxiety. Although Cryan is optimistic that such findings may point the way to the development of psychobiotics, he is wary of hype. "We still need a lot more research into the mechanisms by which gut bacteria interact with the brain," he says.

Charles Schmidt is a recipient of the National Association of Science Writers' Science in Society Journalism Award. His work has appeared in *Science*, *Nature Biotechnology*, *Nature Medicine* and *the Washington Post*.